

# VOLUME FRACTIONS OF THE KINEMATIC “NEAR-CRITICAL” SETS OF THE QUANTUM ENSEMBLE CONTROL LANDSCAPE

JASON DOMINY<sup>1</sup> AND HERSCHEL RABITZ<sup>2,1</sup>

<sup>1</sup> PROGRAM IN APPLIED AND COMPUTATIONAL MATHEMATICS, PRINCETON UNIVERSITY

<sup>2</sup> DEPARTMENT OF CHEMISTRY, PRINCETON UNIVERSITY

**ABSTRACT.** An estimate is derived for the volume fraction of a subset  $C_\epsilon^{\mathcal{P}} = \{U : \|\text{grad } J(U)\| \leq \epsilon\} \subset \text{U}(N)$  in the neighborhood of the critical set  $C^{\mathcal{P}} \simeq \text{U}(\mathbf{n})\mathcal{P}\text{U}(\mathbf{m})$  of the kinematic quantum ensemble control landscape  $J(U) = \text{Tr}(U\rho U^\dagger \mathcal{O})$ , where  $U$  represents the unitary time evolution operator,  $\rho$  is the initial density matrix of the ensemble, and  $\mathcal{O}$  is an observable operator. This estimate is based on the Hilbert-Schmidt geometry for the unitary group and a first-order approximation of  $\|\text{grad } J(U)\|^2$ . An upper bound on these near-critical volumes is conjectured and supported by numerical simulation, leading to an asymptotic analysis as the dimension  $N$  of the quantum system rises in which the volume fractions of these “near-critical” sets decrease to zero as  $N$  increases. This result helps explain the apparent lack of influence exerted by the many saddles of  $J$  over the gradient flow.

## 1. INTRODUCTION

Control landscapes are proving to be valuable for providing insights into quantum optimal control theory [1]. A simplification can be achieved by observing that the dynamical landscape  $\tilde{J} : \mathbb{K} \rightarrow \mathbb{R}$  (i.e. the objective functional) – a map taking a control function as input and producing the value of the observable at some final time  $T$  – can be written as a composition of a kinematic landscape  $J$  and a control→propagator map  $V_T$ . Here the kinematic landscape  $J : \text{U}(N) \rightarrow \mathbb{R}$  is a smooth real-valued function on the unitary group, and  $V_T : \mathbb{K} \rightarrow \text{U}(N)$  is defined implicitly through the Schrödinger equation and returns the unitary time evolution operator at the final time  $T$  for each given input control function. The goal of quantum optimal control is generally to maximize the dynamical landscape  $\tilde{J} = J \circ V_T$ .

The latter basic landscape decomposition has been applied to various quantum control objectives, including state-to-state transitions [2–4], general quantum mechanical observables on an ensemble [5–7], and unitary transformation (quantum gate) preparation [8–10]. This area of research has focussed on identifying and characterizing the critical points of the landscapes, revealing important features of the associated gradient flows. In particular, this work has shown that the kinematic landscapes have no suboptimal extrema that can act as traps preventing the gradient flow from reaching a global optimum.

---

*Date:* January 13, 2013.

*E-mail address:* jdominy@princeton.edu, hrabitz@princeton.edu.

The present paper considers one aspect of the optimal control of a quantum mechanical observable on a  $N$ -level system. Such observables can be defined for an ensemble of initial states (whose collective state is described by a density matrix  $\rho$ ) by the kinematic landscape  $J(U) = \text{Tr}(U\rho U^\dagger \mathcal{O})$ , where the observable is represented by a Hermitian matrix  $\mathcal{O}$ . It is known that, in addition to a single global minimum submanifold and a single maximum submanifold, this kinematic landscape generally contains a large number of saddle submanifolds [7]. The number of saddles depends on the eigenstructure of  $\rho$  and  $\mathcal{O}$  and generally increases with the size  $N$  of the quantum system. By continuity, each saddle submanifold is surrounded by a neighborhood in which the norm of the gradient is small. Any trajectory of the gradient flow entering such a neighborhood will be slowed by this small gradient and reach the global maximum less efficiently than trajectories that never enter such neighborhoods. However, since the gradient flow is not attracted to saddles as it is to local maxima, the magnitude of the impact of these saddles on the overall gradient flow of this landscape is unclear.

With this perspective in mind, we derive estimates and bounds for the effective volume fractions of the critical submanifolds relative to the volume of the unitary group. The term “effective volume fraction” is used here to mean the volume fraction of a region around a given critical submanifold in which the norm of the gradient is small, i.e. less than some  $\epsilon > 0$ . These volume fractions will serve as an indicator of the likelihood of a gradient integral curve falling under the influence of the saddles. This influence can have a profound impact on the efficiency of a gradient ascent algorithm (whether implemented in the laboratory or in numerical simulations), or indeed any optimal control algorithm based on local information.

Additionally, we consider the asymptotic behavior of these volume fractions as the system size  $N$  is allowed to increase. By embedding the original  $\rho$  and  $\mathcal{O}$  in these larger spaces, this analysis addresses the role of truncation to finite  $N$  in defining the characteristics of the gradient flow. In particular, this research will show that the volume fractions tend to zero as  $N \rightarrow \infty$ , implying that the saddles should have a vanishingly small impact on the gradient flow. This conclusion helps to explain the observed behavior of numerical quantum optimal control simulations, which show no evident increase in search effort as  $N$  rises.

The paper is organized as follows. Section 2 reviews the structure of the critical submanifolds of  $J(U) = \text{Tr}(U\rho U^\dagger \mathcal{O})$ . The induced Hilbert-Schmidt measures of these submanifolds are computed in Section 3. The near-critical sets are described in Section 4, yielding estimates of the volumes of these sets in terms of a Haar measure. Upper bounds for these volumes are developed in Section 5 and the asymptotic behavior as  $N \rightarrow \infty$  is considered. The results are summarized Section 6. Four appendices are included that detail the Chevalley lattice of direct products of Lie groups, the Hessian operator of  $J$ , the second fundamental forms of the critical submanifolds of  $J$ , and provide supporting arguments for a key conjecture.

## 2. THE CRITICAL SUBMANIFOLDS OF THE LANDSCAPE

The landscape  $J(U) = \text{Tr}(U\rho U^\dagger \mathcal{O})$  is defined on the unitary group  $U(N)$ , where the density matrix  $\rho$  and the observable  $\mathcal{O}$  are both  $N \times N$  Hermitian matrices. Without loss of generality, we will make the simplifying assumption that  $\rho$  and  $\mathcal{O}$  are both diagonal, with monotonically descending diagonal elements. This may be done since passage from the landscape based on  $\rho$  and  $\mathcal{O}$  to the one based on their diagonalized forms represents merely a translation over the unitary group. Since the metrics considered in this paper are bi-invariant, this translation may be neglected.

It will be fruitful at certain points in the analysis that follows to consider  $U(N)$  (with the geometry induced by the real Hilbert-Schmidt inner product on  $\mathbb{C}^{N \times N}$ ) as a Riemannian globally symmetric space. This will be done by fixing a permutation matrix  $\mathcal{P} \in S_N$  and letting  $\mathcal{G} := U(N) \otimes U(N)$  and  $\mathcal{K}_{\mathcal{P}} := \{(V, \mathcal{P}^\dagger V \mathcal{P}) : V \in U(N)\}$ . Then, under the involutive analytic automorphism  $\sigma(V, W) = (\mathcal{P}W\mathcal{P}^\dagger, \mathcal{P}^\dagger V \mathcal{P})$ ,  $(\mathcal{G}, \mathcal{K}_{\mathcal{P}})$  is a symmetric pair and  $\mathcal{G}/\mathcal{K}_{\mathcal{P}}$  is a Riemannian globally symmetric space under any  $\mathcal{G}$ -invariant metric. The map  $\mathcal{G}/\mathcal{K}_{\mathcal{P}} \rightarrow U(N)$  given by  $(V, W)\mathcal{K}_{\mathcal{P}} \mapsto V\mathcal{P}W^\dagger$  is clearly smooth, as is the section  $U \mapsto (\mathbb{I}, U^\dagger \mathcal{P})$ . Since the projection  $\mathcal{G} \rightarrow \mathcal{G}/\mathcal{K}_{\mathcal{P}}$  is a smooth submersion,  $V\mathcal{P}W^\dagger \mapsto (\mathbb{I}, W\mathcal{P}^\dagger V^\dagger \mathcal{P}) \mapsto (V, W)\mathcal{K}_{\mathcal{P}}$  is a smooth inverse to the map  $(V, W)\mathcal{K}_{\mathcal{P}} \mapsto V\mathcal{P}W^\dagger$ , so this map is a diffeomorphism between  $\mathcal{G}/\mathcal{K}_{\mathcal{P}}$  and  $U(N)$ . The group action of  $\mathcal{G}$  on  $U(N)$  induced by the diffeomorphism is then given by  $(V, W) \cdot U = VUW^\dagger$ , and the real Hilbert-Schmidt metric on  $U(N)$  is  $\mathcal{G}$ -invariant, so  $U(N)$  is a Riemannian globally symmetric space under this structure. The projection  $\Phi_{\mathcal{P}} : \mathcal{G} \rightarrow U(N) \simeq \mathcal{G}/\mathcal{K}_{\mathcal{P}}$  can also be viewed as defining  $\mathcal{G}$  as a fiber bundle with base  $U(N)$  and fiber  $\mathcal{K}_{\mathcal{P}}$ .

Now, let  $\mathcal{H} := U(\mathbf{m}) \oplus U(\mathbf{n}) \subset \mathcal{G}$ , where  $U(\mathbf{m}) = U(m_1) \oplus \cdots \oplus U(m_s)$ ,  $U(\mathbf{n}) = U(n_1) \oplus \cdots \oplus U(n_r)$ , and  $\{m_j\}$  and  $\{n_i\}$  are the multiplicities of the unique eigenvalues of  $\mathcal{O}$  and  $\rho$ , respectively. Then the critical set of  $J(U) = \text{Tr}(U\rho U^\dagger \mathcal{O})$  has been shown to comprise a disjoint union of submanifolds of  $U(N)$ , each of the form  $\Phi_{\mathcal{P}}(\mathcal{H})$  [7], where  $\Phi_{\mathcal{P}}(\mathcal{H}) = U(\mathbf{m})\mathcal{P}U(\mathbf{n})$  is the orbit of the point  $\mathcal{P} \in S_N \subset U(N)$  with respect to the induced action of  $\mathcal{H}$  on  $U(N)$ . So  $\Phi_{\mathcal{P}}(\mathcal{H}) \cong \mathcal{H}/\text{Stab}_{\mathcal{H}}(\mathcal{P})$ , where the stabilizer  $\text{Stab}_{\mathcal{H}}(\mathcal{P})$  is the subgroup of  $\mathcal{H}$  given by

$$\text{Stab}_{\mathcal{H}}(\mathcal{P}) = \mathcal{H} \cap \mathcal{K}_{\mathcal{P}} = \{(V, W) \in U(\mathbf{m}) \oplus U(\mathbf{n}) : \Phi_{\mathcal{P}}(V, W) = V\mathcal{P}W^\dagger = \mathcal{P}\} \quad (1)$$

and is identified with  $U(\mathbf{K}) \cong U(\mathbf{m}) \cap \mathcal{P}U(\mathbf{n})\mathcal{P}^\dagger$  in [7], i.e.  $\text{Stab}(\mathcal{P}) = \{(\zeta, \mathcal{P}^\dagger \zeta \mathcal{P}) : \zeta \in U(\mathbf{m}) \cap \mathcal{P}U(\mathbf{n})\mathcal{P}^\dagger\}$ . Here  $\mathbf{K}$  is the  $r \times s$  “contingency table” corresponding to  $\mathbf{m}$ ,  $\mathbf{n}$ , and  $\mathcal{P}$ . For a given permutation  $\mathcal{P}$ ,  $U(\mathbf{m}) \oplus U(\mathbf{n})$  can be expressed as a fiber bundle [11] as

$$\begin{array}{ccc} U(\mathbf{m}) \cap \mathcal{P}U(\mathbf{n})\mathcal{P}^\dagger & \xrightarrow{\zeta \mapsto (\zeta, \mathcal{P}^\dagger \zeta \mathcal{P})} & U(\mathbf{m}) \oplus U(\mathbf{n}) \\ & & \downarrow \Phi_{\mathcal{P}} \\ & & \text{Orb}(\mathcal{P}) \subset U(N). \end{array}$$

We seek to compute the “volume” (measure) of  $\text{Orb}(\mathcal{P})$  as an embedded submanifold of  $U(N)$ , where  $U(N)$  is given the Riemannian metric induced by the real Hilbert-Schmidt inner product on

$\mathbb{C}^{N \times N}$ . We will see that this problem reduces to one of computing the volumes of the Lie groups  $U(\mathbf{K})$  and  $U(\mathbf{m}) \oplus U(\mathbf{n})$  with respect to certain specific geometries.

**Lemma 1.** *For any permutation matrix  $\hat{\mathcal{P}} \in S_N$ , there exists a  $\mathcal{P} \in S_N \cap \text{Orb}(\hat{\mathcal{P}})$  such that  $U(\mathbf{m}) \cap \mathcal{P}U(\mathbf{n})\mathcal{P}^\dagger = U(\mathbf{K})$  in terms of block-diagonal structure and  $U(\mathbf{n}) \cap \mathcal{P}^\dagger U(\mathbf{m})\mathcal{P} = U(\tilde{\mathbf{K}}) = \mathcal{P}^\dagger U(\mathbf{K})\mathcal{P}$  where  $U(\tilde{\mathbf{K}})$  has essentially the same block structure as  $U(\mathbf{K})$ , but the blocks are reordered.*

*Proof.* For any  $\Sigma \in S_N \cap U(\mathbf{m})$  and  $\Gamma \in S_N \cap U(\mathbf{n})$ ,  $\mathcal{P} = \Sigma\hat{\mathcal{P}}\Gamma^\dagger \in S_N \cap \text{Orb}(\mathcal{P})$ . Then  $U(\mathbf{m}) \cap \mathcal{P}U(\mathbf{n})\mathcal{P}^\dagger = \Sigma^\dagger(U(\mathbf{m}) \cap \hat{\mathcal{P}}U(\mathbf{n})\hat{\mathcal{P}}^\dagger)\Sigma$ , so that  $\Sigma$  may be chosen to reorder the rows and columns of each diagonal block of  $U(\mathbf{m})$  to group the matrix elements of  $U(\mathbf{m}) \cap \hat{\mathcal{P}}U(\mathbf{n})\hat{\mathcal{P}}^\dagger$  according to the block of  $U(\mathbf{n})$  from whence they came, creating sub-blocks within each block of  $U(\mathbf{m})$ . This is exactly the structure desired for  $U(\mathbf{K})$ . Similarly,  $U(\mathbf{n}) \cap \mathcal{P}^\dagger U(\mathbf{m})\mathcal{P} = \Gamma(U(\mathbf{n}) \cap \hat{\mathcal{P}}^\dagger U(\mathbf{m})\hat{\mathcal{P}})$ , so that  $\Gamma$  may be likewise chosen to reorder the elements in each diagonal block of  $U(\mathbf{n})$  according to the originating block of  $U(\mathbf{m})$ , constructing  $U(\tilde{\mathbf{K}})$ .  $\square$

Without loss of generality, the remainder of the paper will use the permutation  $\mathcal{P}$  described in the above lemma, so that  $U(\mathbf{m}) \cap \mathcal{P}U(\mathbf{n})\mathcal{P}^\dagger = U(\mathbf{K})$ .

### 3. THE HILBERT-SCHMIDT MEASURE OF THE CRITICAL SUBMANIFOLDS

We now turn to the problem of computing the volumes of the critical submanifolds. To that end, we will choose geometries for  $U(\mathbf{K})$  and  $U(\mathbf{m}) \oplus U(\mathbf{n})$  such that there is a local isometry between  $U(\mathbf{m}) \oplus U(\mathbf{n})$  and  $\text{Orb}(\mathcal{P}) \oplus U(\mathbf{K})$ . Then it will be shown that the volume of  $\text{Orb}(\mathcal{P})$  is just the quotient of the volumes of  $U(\mathbf{m}) \oplus U(\mathbf{n})$  and  $U(\mathbf{K})$  under these specific geometries.

Let  $\mathcal{A}, \mathcal{B}, \mathcal{C}$  denote the following subspaces of  $\mathfrak{u}(\mathbf{m}) \oplus \mathfrak{u}(\mathbf{n})$ :

$$\mathcal{A} = \{(X, \mathcal{P}^\dagger X \mathcal{P}) : X \in \mathfrak{u}(\mathbf{K})\} \quad (2a)$$

$$\mathcal{B} = \{(X, -\mathcal{P}^\dagger X \mathcal{P}) : X \in \mathfrak{u}(\mathbf{K})\} \quad (2b)$$

$$\mathcal{C} = (\mathcal{A} \oplus \mathcal{B})^\perp = \{(Y, Z) : Y \in \mathfrak{u}(\mathbf{m})/\mathfrak{u}(\mathbf{K}) \text{ and } Z \in \mathfrak{u}(\mathbf{n})/\mathfrak{u}(\tilde{\mathbf{K}})\}. \quad (2c)$$

It may be readily verified that these spaces are mutually orthogonal in the Hilbert-Schmidt geometry and span  $\mathfrak{u}(\mathbf{m}) \oplus \mathfrak{u}(\mathbf{n})$ . We extend them by left translation to form mutually orthogonal distributions over  $U(\mathbf{m}) \oplus U(\mathbf{n})$  that span the tangent space at each point.

**Lemma 2.** *The distributions  $\mathcal{A}$ ,  $\mathcal{B}$ , and  $\mathcal{C}$  are right invariant with respect to  $\text{Stab}(\mathcal{P})$ .*

*Proof.* First observe that for any  $\zeta \in U(\mathbf{K})$ ,  $X \in \mathfrak{u}(\mathbf{K})$  if and only if  $\zeta X \zeta^\dagger \in \mathfrak{u}(\mathbf{K})$ . Likewise  $Y \in \mathfrak{u}(\mathbf{m})/\mathfrak{u}(\mathbf{K})$  if and only if  $\langle Y, X \rangle_{\text{HS}} = 0$  for all  $X \in \mathfrak{u}(\mathbf{K})$ , if and only if  $\langle \zeta Y \zeta^\dagger, \zeta X \zeta^\dagger \rangle_{\text{HS}} = 0$  for all  $X \in \mathfrak{u}(\mathbf{K})$ , if and only if  $\zeta Y \zeta^\dagger \in \mathfrak{u}(\mathbf{m})/\mathfrak{u}(\mathbf{K})$  and similarly for  $Z \in \mathfrak{u}(\mathbf{n})/\mathfrak{u}(\tilde{\mathbf{K}})$  with respect to  $\tilde{\zeta}$ . In other words, the subspaces of  $\mathfrak{u}(\mathbf{m}) \oplus \mathfrak{u}(\mathbf{n})$  given by evaluating the distributions  $\mathcal{A}$ ,  $\mathcal{B}$ , and  $\mathcal{C}$  at

the identity are invariant under the adjoint action of  $\text{Stab}(\mathcal{P})$ . Then, for any  $(\zeta, \mathcal{P}^\dagger \zeta \mathcal{P}) \in \text{Stab}(\mathcal{P})$  and  $(V, W) \in \text{U}(\mathbf{m}) \oplus \text{U}(\mathbf{n})$

$$\mathcal{A}_{(V,W) \cdot (\zeta, \mathcal{P}^\dagger \zeta \mathcal{P})} = \{(V, W) \cdot (\zeta, \mathcal{P}^\dagger \zeta \mathcal{P}) \cdot (X, \mathcal{P}^\dagger X \mathcal{P}) : X \in \mathfrak{u}(\mathbf{K})\} \quad (3a)$$

$$= \{(V, W) \cdot (X, \mathcal{P}^\dagger X \mathcal{P}) \cdot (\zeta, \mathcal{P}^\dagger \zeta \mathcal{P}) : X \in \mathfrak{u}(\mathbf{K})\}, \quad (3b)$$

which is the right translation of  $\mathcal{A}_{(V,W)}$  to  $(V, W) \cdot (\zeta, \mathcal{P}^\dagger \zeta \mathcal{P})$ , so that  $\mathcal{A}$  is right invariant with respect to  $\text{Stab}(\mathcal{P})$ . Similar arguments apply to  $\mathcal{B}$  and  $\mathcal{C}$ .  $\square$

Now, denote by  $\hat{\Phi}_{\mathcal{P}}$  the restriction of  $\Phi_{\mathcal{P}}$  to  $\mathcal{H} = \text{U}(\mathbf{m}) \oplus \text{U}(\mathbf{n})$ , and consider the images of these three distributions  $\mathcal{A}$ ,  $\mathcal{B}$ , and  $\mathcal{C}$  under  $d\hat{\Phi}_{\mathcal{P}}$ . Since for any  $(V, W) \in \text{U}(\mathbf{m}) \oplus \text{U}(\mathbf{n})$  and any  $(\delta V, \delta W) \in T_{(V,W)}\text{U}(\mathbf{m}) \oplus \text{U}(\mathbf{n})$ , there exists  $(X, Y) \in \mathfrak{u}(\mathbf{m}) \oplus \mathfrak{u}(\mathbf{n})$  such that  $(\delta V, \delta W) = (VX, WY)$ ,

$$d_{(V,W)}\hat{\Phi}_{\mathcal{P}}(\delta V, \delta W) = d_{(V,W)}\hat{\Phi}_{\mathcal{P}}(VX, WY) = VX\mathcal{P}W^\dagger - V\mathcal{P}YW^\dagger. \quad (4)$$

Then the kernel of  $d_{(V,W)}\hat{\Phi}_{\mathcal{P}}$  is  $\{(VX, WY) \in T_{(V,W)}\text{U}(\mathbf{m}) \oplus \text{U}(\mathbf{n}) : Y = \mathcal{P}^\dagger X \mathcal{P}\}$ , which is exactly the left-invariant distribution  $\mathcal{A}$  evaluated at  $(V, W)$ . When acting on an element  $(V, W) \cdot (X, -\mathcal{P}^\dagger X \mathcal{P}) \in \mathcal{B}_{(V,W)}$ ,  $d_{(V,W)}\hat{\Phi}_{\mathcal{P}}$  yields  $2VX\mathcal{P}W^\dagger$ , and when acting on an element  $(V, W) \cdot (Y, Z) \in \mathcal{C}_{(V,W)}$ ,  $d_{(V,W)}\hat{\Phi}_{\mathcal{P}}$  yields  $VY\mathcal{P}W^\dagger - V\mathcal{P}ZW^\dagger$ . Then, since  $X \in \mathfrak{u}(\mathbf{K})$  and  $Y - \mathcal{P}Z\mathcal{P}^\dagger \in \mathfrak{u}(N)/\mathfrak{u}(\mathbf{K})$ , the  $d_{(V,W)}\hat{\Phi}_{\mathcal{P}}$  images of  $\mathcal{B}_{(V,W)}$  and  $\mathcal{C}_{(V,W)}$  are orthogonal complements under any bi-invariant metric on  $\text{U}(N)$ , in particular the Hilbert-Schmidt metric. Furthermore, for  $X_1, X_2 \in \mathfrak{u}(\mathbf{K})$ ,  $Y_1, Y_2 \in \mathfrak{u}(\mathbf{m})/\mathfrak{u}(\mathbf{K})$ , and  $Z_1, Z_2 \in \mathfrak{u}(\mathbf{n})/\mathfrak{u}(\hat{\mathbf{K}})$ ,

$$\langle d_{\mathbb{I}}\hat{\Phi}_{\mathcal{P}}(X_1, -\mathcal{P}^\dagger X_1 \mathcal{P}), d_{\mathbb{I}}\hat{\Phi}_{\mathcal{P}}(X_2, -\mathcal{P}^\dagger X_2 \mathcal{P}) \rangle_{\text{HS}} = 2\langle (X_1, -\mathcal{P}^\dagger X_1 \mathcal{P}), (X_2, -\mathcal{P}^\dagger X_2 \mathcal{P}) \rangle_{\text{HS}} \quad (5a)$$

$$\langle d_{\mathbb{I}}\hat{\Phi}_{\mathcal{P}}(Y_1, Z_1), d_{\mathbb{I}}\hat{\Phi}_{\mathcal{P}}(Y_2, Z_2) \rangle_{\text{HS}} = \langle (Y_1, Z_1), (Y_2, Z_2) \rangle_{\text{HS}} \quad (5b)$$

since  $\mathcal{P}Z_1\mathcal{P}^\dagger, \mathcal{P}Z_2\mathcal{P}^\dagger \in (\mathcal{P}\mathfrak{u}(\mathbf{n})\mathcal{P}^\dagger)/\mathfrak{u}(\mathbf{K}) \subset \mathfrak{u}(N)/\mathfrak{u}(\mathbf{m})$  are orthogonal to  $Y_1$  and  $Y_2$  in the Hilbert-Schmidt metric. So the image through  $d\hat{\Phi}_{\mathcal{P}}$  of  $\mathcal{A}$  is zero, the restriction of  $d\hat{\Phi}_{\mathcal{P}}$  to  $\mathcal{B}$  is two times a linear isometry, and the restriction to  $\mathcal{C}$  is a linear isometry.

Since  $\hat{\Phi}_{\mathcal{P}}$  is a fiber bundle with base  $\text{Orb}(\mathcal{P})$  and fiber  $\text{U}(\mathbf{K})$ , for any  $U_0 \in \text{Orb}(\mathcal{P})$ , there exists a neighborhood  $Q$  of  $U_0$  and a smooth local section  $\gamma_Q : Q \rightarrow \text{U}(\mathbf{m}) \oplus \text{U}(\mathbf{n})$  such that  $\hat{\Phi}_{\mathcal{P}}(\gamma_Q(U)) = U$  for all  $U \in Q$  and such that  $\text{Im}(d_U\gamma_Q) = \mathcal{B}_{\gamma_Q(U)} \oplus \mathcal{C}_{\gamma_Q(U)}$ , the orthogonal complement of  $\mathcal{A}_{\gamma_Q(U)} = \ker(d_{\gamma_Q(U)}\hat{\Phi}_{\mathcal{P}})$  in  $T_{\gamma_Q(U)}\text{U}(\mathbf{m}) \oplus \text{U}(\mathbf{n})$  with respect to the Hilbert-Schmidt metric. Let  $\Psi_Q : Q \oplus \text{U}(\mathbf{K}) \rightarrow \hat{\Phi}_{\mathcal{P}}^{-1}(Q) \subset \text{U}(\mathbf{m}) \oplus \text{U}(\mathbf{n})$  be defined by

$$\Psi_Q(U, \zeta) := \gamma_Q(U) \cdot (\zeta, \mathcal{P}^\dagger \zeta \mathcal{P}). \quad (6)$$

Then

$$d_{(U,\zeta)}\Psi_Q(\delta U, \delta \zeta) = \gamma_Q(U) \cdot (\delta \zeta, \mathcal{P}^\dagger \delta \zeta \mathcal{P}) + d_U\gamma_Q(\delta U) \cdot (\zeta, \mathcal{P}^\dagger \zeta \mathcal{P}) \quad (7)$$

where it may be observed that  $\gamma_Q(U) \cdot (\delta \zeta, \mathcal{P}^\dagger \delta \zeta \mathcal{P}) \in \mathcal{A}_{\Psi_Q(U,\zeta)}$  by left invariance and the fact that  $(\delta \zeta, \mathcal{P}^\dagger \delta \zeta \mathcal{P}) \in \mathcal{A}_{(\zeta, \mathcal{P}^\dagger \zeta \mathcal{P})}$  and also  $d_U\gamma_Q(\delta U) \cdot (\zeta, \mathcal{P}^\dagger \zeta \mathcal{P}) \in \mathcal{B}_{\Psi_Q(U,\zeta)} \oplus \mathcal{C}_{\Psi_Q(U,\zeta)}$  by right invariance and the definition of  $\gamma_Q$  whereby  $d_U\gamma_Q(\delta U) \in \mathcal{B}_{\gamma_Q(U)} \oplus \mathcal{C}_{\gamma_Q(U)}$ .

For any given  $\mathcal{P}$ , we now define a Riemannian metric  $\langle \cdot, \cdot \rangle_{\mathcal{P}}$  on  $U(\mathbf{m}) \oplus U(\mathbf{n})$  at the point  $\Upsilon = (V, W)$  as follows. For any  $\delta\Upsilon_1, \delta\Upsilon_2 \in T_{\Upsilon}(U(\mathbf{m}) \oplus U(\mathbf{n}))$ , let

$$\langle \delta\Upsilon_1, \delta\Upsilon_2 \rangle_{\mathcal{P}} := \frac{1}{2} \langle \delta\Upsilon_1^{\mathcal{A}}, \delta\Upsilon_2^{\mathcal{A}} \rangle_{\text{HS}} + \langle d_{\Upsilon} \hat{\Phi}_{\mathcal{P}}(\delta\Upsilon_1^{\mathcal{B}} + \delta\Upsilon_1^{\mathcal{C}}), d_{\Upsilon} \hat{\Phi}_{\mathcal{P}}(\delta\Upsilon_2^{\mathcal{B}} + \delta\Upsilon_2^{\mathcal{C}}) \rangle_{\text{HS}} \quad (8a)$$

$$= \frac{1}{2} \langle \delta\Upsilon_1^{\mathcal{A}}, \delta\Upsilon_2^{\mathcal{A}} \rangle_{\text{HS}} + 2 \langle \delta\Upsilon_1^{\mathcal{B}}, \delta\Upsilon_2^{\mathcal{B}} \rangle_{\text{HS}} + \langle \delta\Upsilon_1^{\mathcal{C}}, \delta\Upsilon_2^{\mathcal{C}} \rangle_{\text{HS}}, \quad (8b)$$

where we have used the orthogonal decomposition  $\delta\Upsilon = \delta\Upsilon^{\mathcal{A}} + \delta\Upsilon^{\mathcal{B}} + \delta\Upsilon^{\mathcal{C}}$  of  $\delta\Upsilon$  into the three orthogonal subspaces given by  $\mathcal{A}$ ,  $\mathcal{B}$ , and  $\mathcal{C}$ . Then for any  $U \in Q$  and  $\zeta \in U(\mathbf{K})$ ,

$$\begin{aligned} & \langle d_{(U, \zeta)} \Psi_Q(\delta U_1, \delta \zeta_1), d_{(U, \zeta)} \Psi_Q(\delta U_2, \delta \zeta_2) \rangle_{\mathcal{P}} \\ &= \frac{1}{2} \langle \gamma_Q(U) \cdot (\delta \zeta_1, \mathcal{P}^{\dagger} \delta \zeta_1 \mathcal{P}), \gamma_Q(U) \cdot (\delta \zeta_2, \mathcal{P}^{\dagger} \delta \zeta_2 \mathcal{P}) \rangle_{\text{HS}} \\ & \quad + \langle d_{\Psi_Q(U, \zeta)} \hat{\Phi}_{\mathcal{P}}(d_U \gamma_Q(\delta U_1) \cdot (\zeta, \mathcal{P}^{\dagger} \zeta \mathcal{P})), d_{\Psi_Q(U, \zeta)} \hat{\Phi}_{\mathcal{P}}(d_U \gamma_Q(\delta U_2) \cdot (\zeta, \mathcal{P}^{\dagger} \zeta \mathcal{P})) \rangle_{\text{HS}} \end{aligned} \quad (9a)$$

$$= \langle \delta \zeta_1, \delta \zeta_2 \rangle_{\text{HS}} + \langle d_{\gamma_Q(U)} \hat{\Phi}_{\mathcal{P}}(d_U \gamma_Q(\delta U_1)), d_{\gamma_Q(U)} \hat{\Phi}_{\mathcal{P}}(d_U \gamma_Q(\delta U_2)) \rangle_{\text{HS}} \quad (9b)$$

$$= \langle \delta \zeta_1, \delta \zeta_2 \rangle_{\text{HS}} + \langle \delta U_1, \delta U_2 \rangle_{\text{HS}}, \quad (9c)$$

where the last step follows from the fact that  $\hat{\Phi}_{\mathcal{P}} \circ \gamma_Q$  is the identity map on  $Q$ . So, if we endow  $U(\mathbf{K}) \subset \mathbb{C}^{N \times N}$  with the Riemannian metric induced by the real Hilbert-Schmidt inner product on  $\mathbb{C}^{N \times N}$ , then  $\Psi_Q$  is an *isometric* diffeomorphism between  $Q \oplus U(\mathbf{K})$  and  $\hat{\Phi}_{\mathcal{P}}^{-1}(Q) \subset U(\mathbf{m}) \oplus U(\mathbf{n})$ . It then follows that  $\text{Vol}_{\text{HS}}(Q) \text{Vol}_{\text{HS}}(U(\mathbf{K})) = \text{Vol}_{\mathcal{P}}(\hat{\Phi}_{\mathcal{P}}^{-1}(Q))$ , and therefore

$$\text{Vol}_{\text{HS}}(\text{Orb}(\mathcal{P})) = \frac{\text{Vol}_{\mathcal{P}}(U(\mathbf{m}) \oplus U(\mathbf{n}))}{\text{Vol}_{\text{HS}}(U(\mathbf{K}))} \quad (10)$$

so the problem reduces to computing the ratio of the volumes of the two Lie groups with respect to the indicated geometries.

The expression in (10) may be simplified further by considering the volume form on  $U(\mathbf{m}) \oplus U(\mathbf{n})$  induced by the metric  $\langle \cdot, \cdot \rangle_{\mathcal{P}}$  [12]. This volume form can be realized by choosing orthonormal bases of  $\mathcal{A}$ ,  $\mathcal{B}$ , and  $\mathcal{C}$  at  $(\mathbb{I}, \mathbb{I})$  in the  $\langle \cdot, \cdot \rangle_{\mathcal{P}}$  metric and extending them to orthonormal vector fields by left translation. Denote these fields by  $\{F_i\}$  where  $i = 1, \dots, d$  where  $d = \sum m_j^2 + \sum n_i^2$ . Then construct the dual basis of 1-forms  $\omega_i = \langle F_i, \cdot \rangle_{\mathcal{P}}$  and the volume form by  $\omega_i \wedge \dots \wedge \omega_d$ . Let  $\kappa = \sum k_{ij}^2 = \|\mathbf{K}\|_{\text{HS}}^2$ . Because of the relationship between  $\langle \cdot, \cdot \rangle_{\mathcal{P}}$  and the Hilbert-Schmidt metric described in (8), if  $F_1, \dots, F_{\kappa}$  is the basis for  $\mathcal{A}$  and  $F_{\kappa+1}, \dots, F_{2\kappa}$  is the basis for  $\mathcal{B}$  under  $\langle \cdot, \cdot \rangle_{\mathcal{P}}$ , then  $\frac{1}{\sqrt{2}}F_1, \dots, \frac{1}{\sqrt{2}}F_{\kappa}$  and  $\sqrt{2}F_{\kappa+1}, \dots, \sqrt{2}F_{2\kappa}$  are the corresponding orthonormal vector fields under the Hilbert-Schmidt metric. So  $\sqrt{2}\omega_1, \dots, \sqrt{2}\omega_{\kappa}$  and  $\frac{1}{\sqrt{2}}\omega_{\kappa+1}, \dots, \frac{1}{\sqrt{2}}\omega_{2\kappa}$  are the corresponding 1-forms under the Hilbert-Schmidt metric. As a result, the volume form  $\hat{\omega}$  on  $U(\mathbf{m}) \oplus U(\mathbf{n})$  induced by the Hilbert-Schmidt metric is identical to  $\omega$ , the volume form induced by  $\langle \cdot, \cdot \rangle_{\mathcal{P}}$ :

$$\hat{\omega} = \sqrt{2}\omega_1 \wedge \dots \wedge \sqrt{2}\omega_{\kappa} \wedge \frac{1}{\sqrt{2}}\omega_{\kappa+1} \wedge \dots \wedge \frac{1}{\sqrt{2}}\omega_{2\kappa} \wedge \omega_{2\kappa+1} \wedge \dots \wedge \omega_d = \omega. \quad (11)$$

So, while  $\langle \cdot, \cdot \rangle_{\mathcal{P}}$  defines a different geometry on  $U(\mathbf{m}) \oplus U(\mathbf{n})$  compared to the Hilbert-Schmidt metric, stretching some dimensions and shrinking others, the result is no net difference in the volume form and therefore no difference in the volume under these two geometries. So we may replace (12) with the expression

$$\text{Vol}_{\text{HS}}(\text{Orb}(\mathcal{P})) = \frac{\text{Vol}_{\text{HS}}(U(\mathbf{m}) \oplus U(\mathbf{n}))}{\text{Vol}_{\text{HS}}(U(\mathbf{K}))}, \quad (12)$$

which reduces the problem to one of computing the volumes of  $U(\mathbf{m}) \oplus U(\mathbf{n})$  and  $U(\mathbf{K})$ , both just direct sums of unitary groups, under the Hilbert-Schmidt metric.

**Lemma 3** (The Volume of  $U(\mathbf{a})$ ). *Let  $\mathbf{a} \in \mathbb{N}^b$  be any vector of non-negative integers. Then the volume of  $U(\mathbf{a}) = U(a_1) \oplus \cdots \oplus U(a_b)$  with respect to the Riemannian metric induced by the real Hilbert-Schmidt inner product is given by*

$$\text{Vol}_{\text{HS}}(U(\mathbf{a})) = \frac{(2\pi)^{\frac{1}{2} \sum a_l^2 + \frac{\bar{a}}{2}}}{\prod_l \prod_{s=0}^{a_l-1} s!}. \quad (13)$$

*Proof.* To apply Macdonald’s formula for the volume of a Lie group [13, 14], we need the basis for the Chevalley lattice given by the vectors  $\{\tau_j\}$ ,  $\{\xi_\alpha\}$ , and  $\{\eta_\alpha\}$  in Appendix A. These vectors are all mutually orthogonal in the real Hilbert-Schmidt inner product on  $U(\mathbf{a})$ . In addition, the  $\tau_j$ ’s have norm 1, and the  $\xi_\alpha$ ’s and  $\eta_\alpha$ ’s have norm  $\sqrt{2}$ . Letting  $\bar{a} = \sum a_l$ , the Gram matrix of this basis is then a diagonal matrix with  $\bar{a}$  entries equal to one and  $\sum_l a_l(a_l - 1)$  entries equal to two, so that the volume of the fundamental cell is just

$$\lambda(\mathfrak{g}/\mathfrak{g}_{\mathbb{Z}}) = 2^{\frac{1}{2} \sum_l a_l(a_l-1)} = 2^{\frac{1}{2} \sum_l a_l^2 - \frac{\bar{a}}{2}}. \quad (14)$$

Then Macdonald’s formula gives the volume of  $U(\mathbf{K})$  as

$$\text{Vol}_{\text{HS}}(U(\mathbf{a})) = 2^{\frac{1}{2} \sum_l a_l^2 - \frac{\bar{a}}{2}} \prod_l \prod_{s=0}^{a_l-1} \text{Vol}(S^{2s+1}) \quad (15a)$$

$$= \frac{(2\pi)^{\frac{1}{2} \sum a_l^2 + \frac{\bar{a}}{2}}}{\prod_l \prod_{s=0}^{a_l-1} s!} \quad (15b)$$

$$= \prod_l \text{Vol}_{\text{HS}}(U(a_l)). \quad (15c)$$

□

With this lemma, we can now compute the volume of the critical submanifold  $\text{Orb}(\mathcal{P})$ :

$$\text{Vol}_{\text{HS}}(\text{Orb}(\mathcal{P})) = \frac{\text{Vol}_{\text{HS}}(\text{U}(\mathbf{m}) \oplus \text{U}(\mathbf{n}))}{\text{Vol}_{\text{HS}}(\text{U}(\mathbf{K}))} = \frac{\prod \text{Vol}_{\text{HS}}(\text{U}(n_i)) \prod \text{Vol}_{\text{HS}}(\text{U}(m_j))}{\prod_{ij} \text{Vol}_{\text{HS}}(\text{U}(k_{ij}))} \quad (16a)$$

$$= \frac{(2\pi)^{\frac{1}{2} \sum n_i^2 + \frac{1}{2} \sum m_j^2 + N}}{(2\pi)^{\frac{1}{2} \sum k_{ij}^2 + \frac{N}{2}}} \frac{\prod_{ij} \prod_{r=0}^{k_{ij}-1} r!}{\prod_i \prod_{p=0}^{n_i-1} p! \prod_j \prod_{q=0}^{m_j-1} q!} \quad (16b)$$

$$= (2\pi)^{\frac{d+N}{2}} \frac{\prod_{ij} \prod_{r=0}^{k_{ij}-1} r!}{\prod_i \prod_{p=0}^{n_i-1} p! \prod_j \prod_{q=0}^{m_j-1} q!}, \quad (16c)$$

where  $d = \sum m_j^2 + \sum n_i^2 - \sum k_{ij}^2$  is the dimension of  $\text{Orb}(\mathcal{P})$ .

### 3.1. Examples.

**Example 1** (Maximum Submanifold of  $P_{i \rightarrow f}$ ). For any two (non-zero) vectors  $|i\rangle$  and  $|f\rangle$  in  $\mathbb{C}^N$ , let  $\rho = \frac{|i\rangle\langle i|}{\langle i|i\rangle}$  and  $\mathcal{O} = \frac{|f\rangle\langle f|}{\langle f|f\rangle}$ . Then  $J(U) = \text{Tr}(U\rho U^\dagger \mathcal{O})$  represents the transition probability from initial state  $|i\rangle$  to final state  $|f\rangle$ . Translating the problem by diagonalizing  $\rho$  and  $\mathcal{O}$  and sorting the eigenvalues in decreasing order, it is found that  $\rho = \mathcal{O}$  has a single "1" in the (1,1) element and zero elsewhere. The  $N^2 - 2N + 2$  dimensional maximum submanifold then corresponds to the identity permutation and yields the following contingency table [7] and volume:

	$m_1 = 1$	$m_2 = N - 1$
$n_1 = 1$	$k_{11} = 1$	$k_{12} = 0$
$n_2 = N - 1$	$k_{21} = 0$	$k_{22} = N - 1$

$$\text{Vol}_{\text{HS}} \text{Orb}(\mathbb{I}) = \frac{(2\pi)^{\frac{1}{2}(N^2 - N + 2)}}{\prod_{p=0}^{N-2} p!}. \quad (17)$$

**Example 2** (Minimum Submanifold of  $P_{i \rightarrow f}$ ). If, in the previous example, a permutation is used that fails to align the non-zero eigenvalues of  $\rho$  and  $\mathcal{O}$ , then  $\text{Orb}(\mathcal{P})$  is the  $N^2 - 2$  dimensional minimum submanifold, with contingency table and volume:

	$m_1 = 1$	$m_2 = N - 1$
$n_1 = 1$	$k_{11} = 0$	$k_{12} = 1$
$n_2 = N - 1$	$k_{21} = 1$	$k_{22} = N - 2$

$$\text{Vol}_{\text{HS}} \text{Orb}(\mathcal{P}) = \frac{(2\pi)^{\frac{1}{2}(N^2 + N - 2)}}{(N - 2)! \prod_{p=0}^{N-2} p!}. \quad (18)$$

**Example 3** (Fully Non-degenerate  $\rho$  and  $\mathcal{O}$ ). In the case where  $\rho$  and  $\mathcal{O}$  are both fully non-degenerate,  $m_1 = \dots = m_N = 1$ ,  $n_1 = \dots = n_N = 1$  and  $\mathbf{K} = \mathcal{P}^\dagger$ . Then for any critical submanifold  $\text{Orb}(\mathcal{P})$ , all of which are  $N$  dimensional tori,

$$\text{Vol}_{\text{HS}} \text{Orb}(\mathcal{P}) = (2\pi)^N. \quad (19)$$

## 4. THE MEASURE OF THE NEAR-CRITICAL SET

In this section, an estimate is derived for the measure of the near-critical set  $C_\epsilon^\mathcal{P} = \{U : \|\text{grad } J(U)\| \leq \epsilon\} \subset \text{U}(N)$  surrounding the critical submanifold  $\text{Orb}(\mathcal{P})$ . To do that, we first approximate the near-critical set by an ellipsoidal tube about  $\text{Orb}(\mathcal{P})$ .



**4.1. Approximating the Set of Interest.** Let  $U \in \text{Orb}(\mathcal{P})$  be a critical point of  $J$ , and let  $X \in \mathfrak{u}(N)$  be such that  $\|X\| = 1$  and  $UX \in (\text{T}_U \text{Crit}(J))^\perp$ , i.e.  $UX \in \text{T}_U \text{U}(N)$  is orthogonal to the null space of  $\text{Hess}_{J,U} : \text{T}_U \text{U}(N) \rightarrow \text{T}_U \text{U}(N)$  (the Hessian operator of  $J$  at the point  $U \in \text{U}(N)$ ) since  $J$  is Morse-Bott, the necessary properties having been established in [7]. Define  $F_X : \mathbb{R} \rightarrow \mathbb{R}$  by

$$F_X(s) = \|\text{grad } J(U \exp(sX))\|^2. \quad (20)$$

Then  $F_X(0) = 0$  and

$$\frac{dF_X}{ds} = 2 \left\langle \text{grad } J(U \exp(sX)), \text{Hess}_{J,U \exp(sX)} (U \exp(sX)X) \right\rangle \quad (21)$$

so that  $(dF_X/ds)(0) = 0$ , since  $U \in \text{Crit}(J)$ . Furthermore,

$$\begin{aligned} \frac{d^2 F_X}{ds^2} &= 2 \left\langle \text{Hess}_{J,U \exp(sX)} (U \exp(sX)X), \text{Hess}_{J,U \exp(sX)} (U \exp(sX)X) \right\rangle \\ &\quad + 2 \left\langle \text{grad } J(U \exp(sX)), \nabla_{U \exp(sX)X} \text{Hess}_{J,U \exp(sX)} (U \exp(sX)X) \right\rangle \end{aligned} \quad (22)$$

so that  $(d^2 F_X/ds^2)(0) = 2\|\text{Hess}_{J,U}(UX)\|^2$ . Thus, for small  $s$ ,  $F_X(s) = s^2\|\text{Hess}_{J,U}(UX)\|^2 + \mathcal{O}(s^3)$ . Then in order to have  $\|\text{grad } J(U \exp(sX))\| \leq \epsilon$ , we should have  $F_X(s) \leq \epsilon^2$ , and therefore  $s^2 \leq \frac{\epsilon^2}{\|\text{Hess}_{J,U}(UX)\|^2} + \mathcal{O}(\epsilon^3)$ .

Now, suppose that  $\{UY_i\}$  are the orthonormal eigenvectors of  $\text{Hess}_{J,U}$  corresponding to non-zero eigenvalues  $\{\beta_i\}$  (see Appendix B). Then each normalized  $UX \in (\text{T}_U \text{Crit}(J))^\perp$  is such that  $X$  can be written  $X = \sum \alpha_i Y_i$  with  $\sum \alpha_i^2 = 1$ . Then the condition on  $s$  is that  $s^2 \leq \frac{\epsilon^2}{\sum \alpha_i^2 \beta_i^2}$ , i.e.  $\sum (s\alpha_i)^2 \beta_i^2 \leq \epsilon^2$ , so that  $(s\alpha_1, \dots, s\alpha_m)$  is a point in the  $m$ -dimensional solid ellipsoid with principal axes  $\{\epsilon/|\beta_i|\}$ . Therefore  $sX$  lies in the  $m$ -dimensional solid ellipsoid with principal axes  $\{(\epsilon/|\beta_i|)Y_i\}$ , and the set of all  $U \exp(sX)$  for  $sX$  in this ellipsoid is an  $m$ -dimensional geodesic ellipsoid in  $\text{U}(N)$ , which we will denote by  $\mathfrak{E}_\epsilon(U)$ . Repeating this analysis at every point  $U$  of the critical submanifold  $\text{Orb}(\mathcal{P})$  and drawing together the resulting geodesic ellipsoids yields an ellipsoidal tube  $\mathfrak{T}_\epsilon = \cup_{U \in \text{Orb}(\mathcal{P})} \mathfrak{E}_\epsilon(U)$  about  $\text{Orb}(\mathcal{P})$  that approximates the set of points for which  $\|\text{grad } J\| \leq \epsilon$ . It is this tube of near-critical points whose volume we will estimate. Figure 1 offers one simple example of such an ellipsoidal tube about the submanifold  $S^1$  of  $\mathbb{R}^3$ .

**4.2. Volumes of Tubes.** The study of the volumes of tubes goes back to 1939, when Hermann Weyl [15] gave the first complete description of the volumes of *spherical* tubes about submanifolds of Euclidean and spherical spaces. These volumes were presented as a finite power series in the radius of tube, with coefficients derived from the geometry of the submanifold. This beautiful result, frequently referred to as Weyl’s tube formula, has been successfully extended to certain other very special spaces (e.g. projective spaces), but appears not to have been extended to Lie groups or to the unitary group in particular. However, we may still make use of an infinite power series [16] to approximate the desired volume.

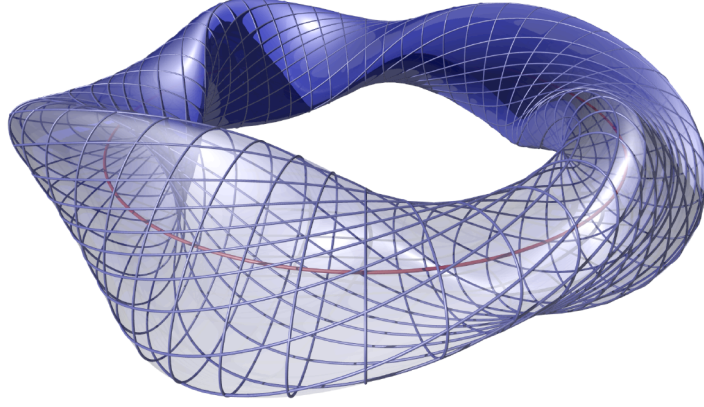


FIGURE 1. Schematic rendering of a simple example of an ellipsoidal tube about a circular 1-dimensional submanifold of  $\mathbb{R}^3$ . The ellipsoidal tubes considered in this paper are of a similar nature, but more complex, being tubes about higher dimensional (and more topologically and geometrically interesting) submanifolds of  $U(N)$ .

We will closely follow the notation, definitions, and conventions in [16]. Let  $\exp_\nu : \nu \rightarrow U(N)$  denote the Riemannian exponential map from the normal bundle  $\nu$  of  $\text{Orb}(\mathcal{P})$  to  $U(N)$ , i.e.  $\exp_\nu$  takes a point  $(U, V) \in \nu$  with  $V \in (T_U \text{Orb}(\mathcal{P}))^\perp$ , and outputs the point in  $U(N)$  found by following the constant speed geodesic  $\xi$  with  $\xi(0) = U$  and  $\xi'(0) = V$  out to  $\xi(1)$ . Let  $y_1, \dots, y_d$  be a local coordinate system for  $\text{Orb}(\mathcal{P})$ . Let  $x_1, \dots, x_{N^2}$  be the Fermi coordinates [16] on  $U(N)$  generated by  $y_1, \dots, y_d$  and the orthonormal fields  $E_{d+1}, \dots, E_{N^2}$  on  $\text{Orb}(\mathcal{P})$  defined as the orthonormal eigenvectors of the Hessian of  $J$  corresponding to non-zero eigenvalues, as in (51),

$$x_i \left( \exp_\nu \left( U, \sum_{j=d+1}^{N^2} t_j E_j(U) \right) \right) = \begin{cases} y_i(U) & 1 \leq i \leq d \\ t_i & d+1 \leq i \leq N^2 \end{cases} \quad (23)$$

Note that the vector fields  $\{E_i\}$  track the principal axes of the ellipsoid normal to  $\text{Orb}(\mathcal{P})$ . For each  $i \in \{1, \dots, N^2\}$ , let  $X_i$  denote the vector field  $\frac{\partial}{\partial x_i}$ . The geodesic ellipsoidal tubular shell characterized by the “radius”  $r$ , denoted  $\tilde{\mathfrak{E}}_r$ , is described in this local coordinate chart by the set of points for which  $\sum_{i=d+1}^{N^2} x_i^2 \beta_i^2 = r^2$ . Let  $L \in \mathfrak{X}(U(N) - \text{Orb}(\mathcal{P}))$  denote the (outward) unit normal vector field to the shells  $\tilde{\mathfrak{E}}_r$ , let  $\hat{L}$  denote the differential 1-form  $\hat{L} = \langle L, \cdot \rangle$ , and let  $\mu$  denote the volume form of  $\tilde{\mathfrak{E}}_r$ . Then  $\hat{L} \wedge \mu = \omega$ , where  $\omega$  is the volume form on  $U(N)$ . Likewise let  $L_\nu$  be the unit normal vector field to the  $\epsilon$ -ellipsoid in  $(T_U \text{Orb}(\mathcal{P}))^\perp$  and let  $\hat{L}_\nu$  be the 1-form  $\hat{L}_\nu = \langle L_\nu, \cdot \rangle$ . Let  $\mu_\nu$  denote the volume form of the ellipsoidal shell in  $(T_U \text{Orb}(\mathcal{P}))$ , so that  $\hat{L}_\nu \wedge \mu_\nu = \omega_\nu$  where  $\omega_\nu$  is the volume form on the normal bundle  $\nu$ .

We would like to compare  $\exp_\nu^*(\hat{L})$ , the pullback of  $\hat{L}$  to  $\nu$ , with  $\hat{L}_\nu$ . To that end, observe that at a point  $V = \sum_{i=d+1}^{N^2} t_i E_i$  in the normal space  $(T_U \text{Orb}(\mathcal{P}))^\perp$ , the tangent space to the ellipsoidal

shell through  $V$  is the set  $\{\sum_{i=d+1}^{N^2} \tau_i E_i\}$  where  $\sum_{i=d+1}^{N^2} \tau_i t_i \beta_i^2 = 0$ . Then the unit normal is  $L_\nu(U, V) = \frac{\sum_{i=d+1}^{N^2} \beta_i^2 t_i E_i}{\sqrt{\sum_{j=d+1}^{N^2} \beta_j^4 t_j^2}} \in (\mathrm{T}_U \mathrm{Orb}(\mathcal{P}))^\perp$ . At the corresponding point  $W = \exp_\nu(U, V) \in \mathrm{U}(N)$ , the tangent space to the geodesic ellipsoid is the set

$$\left\{ \mathrm{d} \exp_\nu \left( \sum_{i=d+1}^{N^2} \tau_i E_i \right) \right\} = \left\{ \sum_{i=d+1}^{N^2} \tau_i X_i \right\} \quad (24)$$

where  $\sum_{i=d+1}^{N^2} \tau_i t_i \beta_i^2 = 0$ . Then  $L(W)$  is the unit vector in  $\mathrm{T}_W \mathrm{U}(N)$  normal to this space, i.e.,  $\sum_{i=d+1}^{N^2} \tau_i \langle L, X_i \rangle = 0$  for all  $\{\tau_i\}$  such that  $\sum_{i=d+1}^{N^2} \tau_i t_i \beta_i^2 = 0$ . This implies that  $\langle L, X_i \rangle = c t_i \beta_i^2$  for all  $i \in \{d+1, \dots, N^2\}$  and some fixed normalization constant  $c$ . Let  $Z \in (\mathrm{T}_U \mathrm{Orb}(\mathcal{P}))^\perp$  be perpendicular to  $L$  [i.e., tangent to the ellipsoid through  $(U, V)$ ]. Then by (24)  $\mathrm{d} \exp_\nu(Z)$  is tangent to the geodesic ellipsoid through  $W$ , so that  $\langle (\mathrm{d} \exp_\nu)^*(L), Z \rangle = \langle L, \mathrm{d} \exp_\nu(Z) \rangle = 0$ , meaning that  $(\mathrm{d} \exp_\nu)^*(L) = c L_\nu$ , where it can be shown that  $c = \|(\mathrm{d} \exp_\nu)^*(L)\| = 1 + O(\epsilon^2)$ . Hence  $\hat{L}_\nu = (1 + O(\epsilon^2)) \exp_\nu^*(\hat{L})$ .

With the above material in mind, and taking  $\mathrm{d}V(U)$  to be the volume measure of the ellipsoid within  $(\mathrm{T}_U \mathrm{Orb}(\mathcal{P}))^\perp$  and  $\mathrm{d}P$  to be the volume measure of  $\mathrm{Orb}(\mathcal{P})$ , we find that the pull-back of  $\omega$  is [16]

$$\begin{aligned} \exp_\nu^*(\hat{L}) \wedge \exp_\nu^*(\mu) &= \exp_\nu^*(\omega)(U, V) = \omega(X_1, \dots, X_{N^2})(\exp_\nu(U, V)) \omega_\nu(U, V) \\ &= \omega(X_1, \dots, X_{N^2})(\exp_\nu(U, V)) L_\nu \wedge \mu_\nu(U, V) \end{aligned} \quad (25a)$$

$$= \omega(X_1, \dots, X_{N^2})(\exp_\nu(U, V)) L_\nu \wedge \mathrm{d}V \wedge \mathrm{d}P \quad (25b)$$

$$= (1 + O(\epsilon^2)) \omega(X_1, \dots, X_{N^2})(\exp_\nu(U, V)) \exp_\nu^*(\hat{L}) \wedge \mathrm{d}V \wedge \mathrm{d}P \quad (25c)$$

and therefore

$$\exp_\nu^*(\mu) = (1 + O(\epsilon^2)) \omega(X_1, \dots, X_{N^2})(\exp_\nu(U, V)) \mathrm{d}V \wedge \mathrm{d}P. \quad (26)$$

Now, it was shown in [16] that

$$\omega(X_1, \dots, X_{N^2}) = 1 - \sum_{i=d+1}^{N^2} \langle H, X_i \rangle x_i + \text{higher order terms}, \quad (27)$$

where  $H$  is a section of the normal bundle  $\nu$  over  $\mathrm{Orb}(\mathcal{P})$  called the mean curvature field. The first order term in the above expression for  $\omega$  will integrate to zero (as will all odd order terms) due to

the symmetry of the ellipse. Then, the area of the geodesic ellipsoidal shell  $\tilde{\mathfrak{E}}_\epsilon$  is given by

$$\text{Area}(\tilde{\mathfrak{E}}_\epsilon) = \int_{\tilde{\mathfrak{E}}_\epsilon} d\mu = \int_{\mathfrak{E}_\epsilon} \exp_\nu^*(d\mu) \quad (28a)$$

$$= \int_{\text{Orb}(\mathcal{P})} \int_{\text{Ellipse}(\epsilon)} (1 + O(\epsilon^2)) \omega(X_1, \dots, X_{N^2}) dV \wedge dP \quad (28b)$$

$$= (1 + O(\epsilon^2)) \epsilon^{N^2-d-1} \int_{\text{Orb}(\mathcal{P})} \int_{\text{Ellipse}(1)} \omega(X_1, \dots, X_{N^2}) dV \wedge dP \quad (28c)$$

$$= \epsilon^{N^2-d-1} \text{Vol}_{\text{HS}}(\text{Orb}(\mathcal{P})) \text{Vol}(\text{Ellipse}(1)) + O(\epsilon^{N^2-d+1}) \quad (28d)$$

where  $\text{Ellipse}(r)$  is the ellipse with principal axes  $r/|\beta_i|$ .

So we conclude that for small enough  $\epsilon > 0$ , the volume of the ellipsoidal tube about  $\text{Orb}(\mathcal{P})$  is

$$\text{Vol}(\mathfrak{T}_\epsilon) = \int_0^\epsilon \text{Area}(\tilde{\mathfrak{E}}_r) dr \quad (29a)$$

$$= \frac{\epsilon^{N^2-d}}{N^2-d} \text{Vol}_{\text{HS}}(\text{Orb}(\mathcal{P})) \text{Vol}(\text{Ellipse}(1)) + O(\epsilon^{N^2-d+2}) \quad (29b)$$

$$= \frac{2^{\frac{d+N}{2}} \pi^{\frac{N(N+1)}{2}} \epsilon^{N^2-d}}{\Gamma(\frac{N^2-d}{2} + 1) \prod |\beta_i|} \times \frac{\prod_{ij} \prod_{r=0}^{k_{ij}-1} r!}{\prod_i \prod_{p=0}^{n_i-1} p! \prod_j \prod_{q=0}^{m_j-1} q!} + O(\epsilon^{N^2-d+2}), \quad (29c)$$

where  $d = \sum n_i^2 + \sum m_j^2 \sum k_{ij}^2 = \dim(\text{Orb}(\mathcal{P}))$ . Then the volume fraction of the tube within  $U(N)$  is

$$\text{VolFrac}(\mathfrak{T}_\epsilon) = \frac{\text{Vol}(\mathfrak{T}_\epsilon)}{\text{Vol}_{\text{HS}}(U(N))} = \frac{\prod_{s=0}^{N-1} s!}{(2\pi)^{\frac{N(N+1)}{2}}} \text{Vol}(\mathfrak{T}_\epsilon) \quad (30a)$$

$$= \frac{\epsilon^{N^2-d}}{2^{\frac{N^2-d}{2}} \Gamma(\frac{N^2-d}{2} + 1) \prod |\beta_i|} \times \frac{\prod_{s=0}^{N-1} s! \prod_{ij} \prod_{r=0}^{k_{ij}-1} r!}{\prod_i \prod_{p=0}^{n_i-1} p! \prod_j \prod_{q=0}^{m_j-1} q!} + O(\epsilon^{N^2-d+2}). \quad (30b)$$

**4.3. Examples.** We now return to the examples considered in Section 3.1 and compute the volume fractions of the corresponding near-critical sets.

**Example 4** (Maximum Submanifold of  $P_{i \rightarrow f}$ ). *The Hessian of  $P_{i \rightarrow f}$  on the maximum submanifold has rank  $2N - 2$ , and all of the nonzero eigenvalues are  $\beta_i = -1$ . Then, the volume of  $\text{Ellipse}(1)$  with principal axes  $1/|\beta_i|$  is just the volume of the unit sphere  $S^{2N-2}$ , which is  $(2N - 1) \frac{2^N \pi^{N-1}}{(2N-1)!!} = \frac{2^{2N-1} \pi^{N-1} (N-1)!}{(2N-2)!}$  (the double factorial  $(2N - 1)!!$  is defined as the product of the odd integers from*

1 to  $2N - 1$ ). Using the volume computed in Section 3.1 for  $\text{Orb}(\mathbb{I})$ , we find

$$\text{VolFrac}(\mathfrak{T}_\epsilon) = \frac{\epsilon^{2N-2}}{2N-2} \frac{\text{Vol}_{\text{HS}} \text{Orb}(\mathbb{I}) \text{Vol}(\text{Ellipse}(1))}{\text{Vol}_{\text{HS}}(\text{U}(N))} + O(\epsilon^{2N}) \quad (31a)$$

$$= \frac{\epsilon^{2N-2}}{2N-2} \frac{(2\pi)^{\frac{1}{2}(N^2-N+2)}}{\prod_{p=0}^{N-2} p!} \frac{2^{2N-1} \pi^{N-1} (N-1)!}{(2N-2)!} \frac{\prod_{s=0}^{N-1} s!}{(2\pi)^{\frac{N(N+1)}{2}}} + O(\epsilon^{2N}) \quad (31b)$$

$$= \epsilon^{2N-2} \frac{2^{N-1} (N-1)! (N-2)!}{(2N-2)!} + O(\epsilon^{2N}) \quad (31c)$$

$$= \epsilon^{2N-2} \frac{(N-2)!}{(2N-3)!!} + O(\epsilon^{2N}), \quad (31d)$$

where

$$\left(\frac{1}{2}\right)^{N-3} \frac{1}{2N-3} < \frac{(N-2)!}{(2N-3)!!} \leq \left(\frac{2}{3}\right)^{N-3} \frac{1}{2N-3} \quad (32)$$

for  $N > 2$ .

**Example 5** (Minimum Submanifold of  $P_{i \rightarrow f}$ ). The Hessian of  $P_{i \rightarrow f}$  on the minimum submanifold has rank 2, with both nonzero eigenvalues equal to one. So the volume of  $\text{Ellipse}(1)$  is just the volume of the unit sphere  $S^2$ , which is  $4\pi$ . Using the volume computed in Section 3.1 for  $\text{Orb}(\mathcal{P})$ , we find

$$\text{VolFrac}(\mathfrak{T}_\epsilon) = \frac{\epsilon^2}{2} \frac{\text{Vol}_{\text{HS}} \text{Orb}(\mathcal{P}) \text{Vol}(\text{Ellipse}(1))}{\text{Vol}_{\text{HS}}(\text{U}(N))} + O(\epsilon^4) \quad (33a)$$

$$= \frac{\epsilon^2}{2} \frac{(2\pi)^{\frac{1}{2}(N^2+N-2)}}{(N-2)! \prod_{p=0}^{N-2} p!} (4\pi) \frac{\prod_{s=0}^{N-1} s!}{(2\pi)^{\frac{N(N+1)}{2}}} + O(\epsilon^4) \quad (33b)$$

$$= (N-1)\epsilon^2 + O(\epsilon^4). \quad (33c)$$

**Example 6** (Fully Non-degenerate  $\rho$  and  $\mathcal{O}$ ). In the case where  $\rho$  and  $\mathcal{O}$  are both fully non-degenerate,  $m_1 = \dots = m_N = 1$ ,  $n_1 = \dots = n_N = 1$  and  $\mathbf{K} = \mathcal{P}^\dagger$ . Then for any critical submanifold  $\text{Orb}(\mathcal{P})$ , all of which are  $N$  dimensional tori,

$$\text{VolFrac}(\mathfrak{T}_\epsilon) = \frac{\epsilon^{N^2-N}}{N^2-N} \frac{\text{Vol}_{\text{HS}} \text{Orb}(\mathcal{P}) \text{Vol}(\text{Ellipse}(1))}{\text{Vol}_{\text{HS}}(\text{U}(N))} + O(\epsilon^{N^2-N+2}) \quad (34a)$$

$$= \frac{\epsilon^{N^2-N} \pi^{\frac{N^2-N}{2}} (2\pi)^N}{\left(\frac{N^2-N}{2}\right)! \prod |\beta_i|} \frac{\prod_{s=0}^{N-1} s!}{(2\pi)^{\frac{N(N+1)}{2}}} + O(\epsilon^{N^2-N+2}) \quad (34b)$$

$$= \frac{\prod_{s=0}^{N-1} s!}{2^{\frac{N(N-1)}{2}} \left(\frac{N^2-N}{2}\right)! \prod |\beta_i|} \epsilon^{N^2-N} + O(\epsilon^{N^2-N+2}) \quad (34c)$$

where  $\prod |\beta_i|$  is the product of the  $N^2 - N$  nonzero Hessian eigenvalues, and where

$$\frac{1}{\prod_{s=1}^{N-1} (s^2 - s + 2)^s} \leq \frac{\prod_{s=0}^{N-1} s!}{2^{\frac{N(N-1)}{2}} \left(\frac{N^2-N}{2}\right)!} \leq \frac{1}{\prod_{s=1}^{N-1} (s+1)^s} = \prod_{s=1}^{N-1} \frac{s!}{N!}. \quad (35)$$

## 5. ASYMPTOTIC ANALYSIS

The expression in (30b) provides a means of estimating the volume fraction of any given near-critical set. However, for probing the general asymptotic behavior of these volumes as the dimension  $N$  of the state space rises, this estimate is inadequate since it only holds for small enough  $\epsilon > 0$ , where “small enough” is determined for each system and each dimension  $N$ . Before seeking a new expression of practical utility, it is necessary to define the parameters of the desired asymptotic analysis. Fix some  $N_0 > 0$ ,  $N_0 \times N_0$  density matrix  $\rho^0$ , and  $N_0 \times N_0$  Hermitian observable operator  $\mathcal{O}^0$ . Then for any  $z \in \mathbb{N}$ , let  $N_z = N_0 + z$ ,  $\rho^z = \rho^0 \oplus 0_z$ , and  $\mathcal{O}^z = \mathcal{O}^0 \oplus 0_z$ , where  $0_z$  is the  $z \times z$  zero matrix. Then each critical value of the kinematic landscape  $J_z(U) = \text{Tr}(U \rho^z U^\dagger \mathcal{O}^z)$  is also a critical value of  $J_{z+1}$ , so each critical submanifold of  $J_z$  has a direct analog in  $J_{z+1}$ . In this fashion, one can describe an infinite sequence of critical submanifolds as  $z \rightarrow \infty$  and consider the asymptotic behavior of the volume fractions of the near-critical sets around these critical submanifolds. It will be argued that these volume fractions converge to zero as  $z \rightarrow \infty$ . The landscape  $J_z$  for  $z > N_0$  has the same number of critical submanifolds as  $J_{N_0}$ , so if the volume fractions of the individual near-critical sets converge to zero, then the total near-critical volume fraction of  $J_z$  also converges to zero as  $z \rightarrow \infty$ .

Because compact Lie groups with bi-invariant metrics have non-negative sectional curvature [17], the following comparison theorem proved in [16, Ch. 8] may be used to bound the volume of spherical tubes about a submanifold  $P$ .

**Theorem 1.** *Let  $M$  be an  $n$ -dimensional Riemannian manifold with non-negative sectional curvature. Then for any  $d$ -dimensional submanifold  $P \subset M$  and all  $r \geq 0$ , the volume of the spherical tube of radius  $r$  about  $P$  in  $M$  is bounded as*

$$V_P^M(r) \leq \int_0^r \int_P \int_{S^{n-d-1}} t^{n-d-1} \max \left( \left( 1 - \frac{t}{d} \langle H, u \rangle \right)^d, 0 \right) du \, dP \, dt, \quad (36)$$

where  $H$  is the mean curvature vector field.

Since the submanifolds of  $U(N)$  considered in this paper are all minimal ( $H = 0$ ), this bound reduces to

$$V_{\text{Orb}(\mathcal{P})}^{U(N)}(r) \leq \frac{r^{N^2-d}}{N^2-d} \text{Vol}(S^{N^2-d-1}) \text{Vol}(\text{Orb}(\mathcal{P})). \quad (37)$$

Of course, any ellipsoidal tube with longest principal axis  $\epsilon/|\beta_{\min}|$  is contained in the spherical tube with radius  $r = \epsilon/|\beta_{\min}|$ . In the sequence of density matrices and observable operators  $\rho^z$  and  $\mathcal{O}^z$  described above, once  $z > N_0$ , the set of distinct Hessian eigenvectors for a given critical

submanifold does not change with  $z$ . Then  $\beta_{\min}$  is fixed and the  $\epsilon$ -ellipsoidal tube about the critical submanifold is contained in the spherical tube with radius  $r = \epsilon/|\beta_{\min}|$ , so the volume of the  $\epsilon$ -ellipsoidal tube is bounded by

$$\text{Vol}(\mathfrak{T}_\epsilon) \leq \frac{\epsilon^{N^2-d}}{(N^2-d)|\beta_{\min}|^{N^2-d}} \text{Vol}(S^{N^2-d-1}) \text{Vol}(\text{Orb}(\mathcal{P})), \quad (38)$$

so that

$$\text{VolFrac}(\mathfrak{T}_\epsilon) \leq \frac{\epsilon^{N^2-d}}{(N^2-d)|\beta_{\min}|^{N^2-d}} \frac{\text{Vol}(S^{N^2-d-1}) \text{Vol}(\text{Orb}(\mathcal{P}))}{\text{Vol}(\text{U}(N))} \quad (39a)$$

$$= \frac{\epsilon^{N^2-d}}{2^{(N^2-d)/2} \left(\frac{N^2-d}{2}\right)! |\beta_{\min}|^{N^2-d}} \frac{\prod_{s=0}^{N-1} s! \prod_{ij} \prod_{r=0}^{k_{ij}-1} r!}{\prod_i \prod_{p=0}^{n_i-1} p! \prod_j \prod_{q=0}^{m_j-1} q!} \quad (39b)$$

where  $d = \sum m_i^2 + \sum n_j^2 - \sum k_{ij}^2$  is the dimension of the critical submanifold.

Now, for  $z > N_0$ , the set of critical values does not change with  $z$ , so fix a critical value  $v$  and consider the contingency table associated with the corresponding critical submanifold. As  $z$  increases, the only elements in the table that will change are  $m_s$ , the multiplicity of the zero eigenvalue of  $\rho$ ,  $n_r$ , the multiplicity of the zero eigenvalue of  $\mathcal{O}$ , and  $k_{sr}$ , the degree of overlap between the zero eigenvalues of  $\Lambda$  and  $\mathcal{P}^\dagger \Sigma \mathcal{P}$ , where  $\Lambda = \Omega \rho \Omega^\dagger$  and  $\Sigma = \Gamma \mathcal{O} \Gamma^\dagger$  are diagonalizations of  $\rho$  and  $\mathcal{O}$  with decreasing elements. For  $z > N_0$ , these indices change as  $m_s(z+1) = m_s(z) + 1$ ,  $n_r(z+1) = n_r(z) + 1$ , and  $k_{sr}(z+1) = k_{sr}(z) + 1$ . Then  $d(z+1) = d(z) + 2(m_s(z) + n_r(z) - k_{sr}(z)) + 1 = d(z) + 2(m_s(N_0) + n_r(N_0) - k_{sr}(N_0)) + 2(z - N_0) + 1$ . So  $N_{z+1}^2 - d(z+1) = N_z^2 - d(z) + 2(2N_0 - m_s(N_0) - n_r(N_0) + k_{sr}(N_0))$ . Let  $D^z(\epsilon)$  denote the right-hand side of (39b) for  $z > N_0$ . Then let

$$F^z(\epsilon) = \frac{D^{z+1}(\epsilon)}{D^z(\epsilon)} = \frac{\epsilon^{2\zeta}}{2^\zeta |\beta_{\min}|^{2\zeta}} \frac{\left(\frac{N_z^2 - d(z)}{2}\right)!}{\left(\frac{N_z^2 - d(z)}{2} + \zeta\right)!} \frac{N_z! k_{sr}(z)!}{n_r(z)! m_s(z)!} \quad (40a)$$

$$= \frac{\epsilon^{2\zeta}}{2^\zeta |\beta_{\min}|^{2\zeta}} \frac{\left(\frac{4N_0^2 - d(N_0)}{2} + (z - N_0)\zeta\right)!}{\left(\frac{4N_0^2 - d(N_0)}{2} + (z + 1 - N_0)\zeta\right)!} \frac{N_z! k_{sr}(z)!}{n_r(z)! m_s(z)!} \quad (40b)$$

where  $\zeta = 2N_0 - m_s(N_0) - n_r(N_0) + k_{sr}(N_0) \geq 0$  since if  $\nu_\rho$  and  $\nu_{\mathcal{O}}$  are the nullities of  $\rho^0$  and  $\mathcal{O}^0$ , then  $m_s(N_0) = N_0 + \nu_\rho$ ,  $n_r(N_0) = N_0 + \nu_{\mathcal{O}}$ , and  $k_{sr}(N_0) \geq \nu_\rho + \nu_{\mathcal{O}}$ . Moreover, note that if  $\zeta = 0$ , then  $k_{sr}(N_0) = \nu_\rho + \nu_{\mathcal{O}}$  and the critical value under consideration must be  $v = 0$ , so for any critical value  $v \neq 0$ ,  $\zeta > 0$ . Let

$$G^z = \frac{F^z(\epsilon)}{F^{z-1}(\epsilon)} = \frac{\left(\frac{4N_0^2 - d(N_0)}{2} + (z - 1 - N_0)\zeta\right)!}{\left(\frac{4N_0^2 - d(N_0)}{2} + (z + 1 - N_0)\zeta\right)!} \frac{(N_0 + z)(k_{sr}(N_0) - N_0 + z)}{(n_r(N_0) - N_0 + z)(m_s(N_0) - N_0 + z)}, \quad (41)$$

where

$$\frac{(N_0 + z)(k_{sr}(N_0) - N_0 + z)}{n_r(N_0) - N_0 + z)(m_s(N_0) - N_0 + z)} = 1 + \zeta \frac{1}{z} + O(z^{-2}). \quad (42)$$



Then whenever  $\zeta > 0$ , the expression in (42) converges to 1 for large  $z$ , so that  $G^z \searrow 0$  as  $z \rightarrow \infty$ . Therefore  $F^z(\epsilon) \searrow 0$  and consequently  $D^z(\epsilon) \searrow 0$  as  $z \rightarrow \infty$ . Moreover, it may be seen that  $G^z = O(z^{-2\zeta})$ , so that  $F^z(\epsilon) = O((z!)^{-2\zeta})$  and  $D^z(\epsilon) = O(\prod_{s=1}^{z-1} (s!)^{-2\zeta})$ . So the volume fraction of the spherical tube of radius  $\epsilon/|\beta_{\min}|$  converges to zero as  $z \rightarrow \infty$  and finally we may conclude that the volume fraction of the ellipsoidal tube approximating the  $\epsilon$  near-critical set about a critical submanifold with critical value  $v \neq 0$  also converges to zero as  $z \rightarrow \infty$ . Since this convergence proceeds very quickly as the negative power of a product of factorials, it is independent of the slower exponential contribution from  $\epsilon^{N^2-d}$ , and therefore independent of the choice of  $\epsilon$ .

This last result demonstrates the convergence of the volume fractions of the approximating ellipsoidal tubes. To improve on this and show the convergence of the volume fractions of the near-critical sets themselves, it will be necessary to make use of the following conjecture. Evidence in support of the conjecture is presented in Appendix D.

**Conjecture 1.** *Let  $U \in \mathbf{U}(N)$  be a critical point of  $J(U) = \text{Tr}(U\rho U^\dagger \mathcal{O})$ , and let  $A \in \mathfrak{u}(\mathfrak{n})$  be such that the tangent vector  $UA \in \mathbf{T}_U \mathbf{U}(N)$  is of unit length and lies normal to the critical submanifold through  $U$ . Define  $f : [0, \pi/(2\sqrt{2})] \rightarrow \mathbb{R}$  to be the norm squared of the gradient of  $J$  along the unit speed geodesic in the direction  $UA$ , i.e.*

$$f(s) = \|\text{grad } J(U \exp(sA))\|^2 = \|\exp(-sA)U^\dagger \mathcal{O} U \exp(sA), \rho\|^2 = \|[U^\dagger \mathcal{O} U, \exp(sA)\rho \exp(-sA)]\|^2, \quad (43)$$

*and let  $\beta_{\min}$  be the minimum (in absolute value) nonzero eigenvalue of the Hessian of  $J$  at  $U$ . Then  $f(s) \geq \beta_{\min}^2 \sin^2(\sqrt{2}s)/2$  for all  $0 \leq s \leq \pi/(2\sqrt{2})$ .*

With this conjecture, once  $z > N_0$ , the set of Hessian eigenvalues no longer changes with  $z$ , so there is one fixed  $\beta_{\min}$  for all  $z > N_0$ . Then for any  $\epsilon < |\beta_{\min}|/\sqrt{2}$ , when  $r = \epsilon\pi/(2|\beta_{\min}|) < \pi/(2\sqrt{2})$ , it is found that  $f(r) \geq \beta_{\min}^2 \sin^2(\sqrt{2}r)/2 \geq 4\beta_{\min}^2 r^2/\pi^2 = \epsilon^2$ , so that the  $\epsilon$  near-critical set about the critical submanifold is contained within the radius  $r$  spherical tube. For any fixed radius, the volume fraction of this spherical tube was shown to converge to zero as  $z \rightarrow \infty$  when the critical value  $v \neq 0$ . So, it is seen that if Conjecture 1 holds, then the volume fraction converges to zero (as the negative power of a product of factorials) for the  $\epsilon$  near-critical set about any critical submanifold with critical value  $v \neq 0$ . Since saddles are non-attractive critical points, this suggests that as  $z$  gets large, the probability becomes vanishingly small that the gradient flow from a (Haar distributed) random point passes through one of these flat “near-critical” regions around a saddle submanifold.

Referring back to the examples from Sections 3.1 and 4.3, it may be seen that, since the transition probability  $P_{i \rightarrow f}$  involves a rank one density matrix  $\rho$  and observable operator  $\mathcal{O}$ , there exist sequences of these landscapes (for increasing  $N$ ) that fit the required behavior for  $\rho^z$  and  $\mathcal{O}^z$  needed for the analysis in this section. Of the two critical submanifolds of this landscape, only the maximum submanifold  $P_{i \rightarrow f} = 1$  satisfies the further condition that  $\zeta$  be nonzero. The remaining example of non-degenerate  $\rho$  and  $\mathcal{O}$  for every  $N$  does not adhere to these requirements and falls outside this analysis. Ultimately, the asymptotic analysis of such an example in which  $\rho^N$  and  $\mathcal{O}^N$



are fully non-degenerate for all  $N$  would be sensitive to the asymptotic behavior of the eigenvalues of  $\rho^N$  and  $\mathcal{O}^N$ .

## 6. CONCLUSIONS

This work computed the volumes of the critical submanifolds of  $J(U) = \text{Tr}(U\rho U^\dagger \mathcal{O})$  in the induced Hilbert-Schmidt measure, and developed estimates and bounds for the volume fractions of near-critical sets of the form  $C_\epsilon^\mathcal{P} = \{U : \|\text{grad } J(U)\| \leq \epsilon\} \subset \text{U}(N)$ . An asymptotic analysis of these volume fractions revealed that, when the critical value is non-zero, the volume fraction converges to zero as  $N \rightarrow \infty$ . This result helps to explain previous observations that numerical quantum optimal control experiments seem not to be adversely affected by the presence of a large number of high-dimensional saddle submanifolds.

The work presented here focussed on the geometry of the kinematic landscape  $J : \text{U}(N) \rightarrow \mathbb{R}$ . Although it is outside the scope of this paper, to relate this work more closely to numerical and laboratory quantum optimal control experiments, these results should be pulled back to the corresponding dynamical landscape  $\tilde{J} : \mathbb{K} \rightarrow \mathbb{R}$  defined on the space of controls. This effort must address a number of difficulties including the dependence of these landscapes on the details of the quantum system and, depending on definitions, dependence on the final time  $T$ . But perhaps the biggest problem is that, in order to make mathematical sense of the concept of volume fractions, a probability measure needs to be defined either explicitly or implicitly on the control space  $\mathbb{K}$ , which is typically infinite dimensional and unbounded, such as  $\mathbb{K} = L^2([0, T]; \mathbb{R})$ . Overcoming these difficulties could provide a clearer picture of the gradient flow of  $\tilde{J}$  on  $\mathbb{K}$ .

## 7. ACKNOWLEDGMENTS

This work was supported, in part, by U.S. Department of Energy (DOE) Contract No. DE-AC02-76-CHO-3073 through the Program in Plasma Science and Technology at Princeton. We also acknowledge support from Lockheed Martin and from DOE grant No. DE-FG02-02ER15344

## APPENDIX A. THE CHEVALLEY LATTICE OF $\text{U}(a_1) \oplus \cdots \oplus \text{U}(a_b)$

The first step to evaluating the volume of a compact Lie group  $G$  via Macdonald’s formula [13, 14] is to work out its Chevalley lattice  $\mathfrak{g}_{\mathbb{Z}}$  (closely related to the concept of the Chevalley basis [18]). In this appendix, we describe this lattice for groups of the form  $G = \text{U}(a_1) \oplus \cdots \oplus \text{U}(a_b)$  as we will need to compute volumes of two such Lie groups to obtain the desired structure of the critical set  $\text{Orb}(\mathcal{P})$ . To that end, first observe that the set of all diagonal matrices in  $G$  forms a maximal abelian torus  $T$ , the Lie algebra of which,  $\mathfrak{t}$ , is the set of all diagonal matrices in  $\mathfrak{g} = \mathfrak{u}(a_1) \oplus \cdots \oplus \mathfrak{u}(a_b)$ . Then defining  $\mathfrak{t}_{\mathbb{Z}}$  such that  $2\pi\mathfrak{t}_{\mathbb{Z}}$  is the kernel of  $\exp : \mathfrak{t} \rightarrow T$ , it is found that  $\mathfrak{t}_{\mathbb{Z}} = i \text{diag}(\mathbb{Z}^{\bar{a}})$  is the lattice of diagonal  $\bar{a} \times \bar{a}$  matrices with imaginary integer elements, where  $\bar{a} = \sum a_j$ .

Now, each  $\mathfrak{u}(a_l)$  has complexification  $\mathfrak{gl}(a_l)$  which is the direct sum of  $\mathfrak{sl}(a_l)$  and a one-dimensional abelian algebra. So the root vectors of  $\mathfrak{gl}(a_l)$  are just those of  $\mathfrak{sl}(a_l)$ . The root space decomposition of  $\mathfrak{sl}(a_l)$  is defined by the positive roots  $\alpha_{jk}^l$  ( $1 \leq j < k \leq a_l$ ) and the corresponding coroots  $H_{\alpha_{jk}^l}$  and root vectors  $X_{\alpha_{jk}^l}$  [19] given by

$$\alpha_{jk}^l(H) = H_{jj} - H_{kk} \quad 1 \leq j < k \leq a_l \quad (44a)$$

$$H_{\alpha_{jk}^l} = |j\rangle\langle j| - |k\rangle\langle k| \quad (44b)$$

$$X_{\alpha_{jk}^l} = |j\rangle\langle k| \quad (44c)$$

$$X_{-\alpha_{jk}^l} = -|k\rangle\langle j|. \quad (44d)$$

where  $H$  is an arbitrary element of the Cartan subalgebra of diagonal matrices in  $\mathfrak{sl}(a_l)$ . If, for any  $X = \xi + i\eta \in \mathfrak{sl}(a_l)$  with  $\xi, \eta \in \mathfrak{su}(a_l)$ , we define  $\tilde{X} = \xi - i\eta$ , then the  $X_\alpha$ 's above satisfy  $X_{-\alpha} = \tilde{X}_\alpha$  as indicated in [13]. Then finally let

$$\xi_{\alpha_{jk}^l} := (\underbrace{0, \dots, 0}_{l-1 \text{ zeros}}, X_{\alpha_{jk}^l} + X_{-\alpha_{jk}^l}, \underbrace{0, \dots, 0}_{b-l \text{ zeros}}) = (0, \dots, 0, |j\rangle\langle k| - |k\rangle\langle j|, 0, \dots, 0) \quad (45a)$$

$$\eta_{\alpha_{jk}^l} := (0, \dots, 0, i(X_{\alpha_{jk}^l} - X_{-\alpha_{jk}^l}), 0, \dots, 0) = (0, \dots, 0, i(|j\rangle\langle k| + |k\rangle\langle j|), 0, \dots, 0) \quad (45b)$$

for all  $1 \leq l \leq b$  and all  $1 \leq j < k \leq a_l$ . These vectors, along with the basis of  $\mathfrak{t}_{\mathbb{Z}}$  given by  $\{\tau_j := i|j\rangle\langle j|\}$  for  $j = 1, \dots, \bar{a}$ , form the basis for the Chevalley lattice of  $\mathfrak{U}(a_1) \oplus \dots \oplus \mathfrak{U}(a_b)$  denoted by  $\mathfrak{g}_{\mathbb{Z}}$  in [13]. The volume of the fundamental cell  $\mathfrak{g}/\mathfrak{g}_{\mathbb{Z}}$  with respect to a given inner product is just the square root of the determinant of the Gram matrix constructed from the basis for  $\mathfrak{g}_{\mathbb{Z}}$ .

## APPENDIX B. THE HESSIAN OF $J(U) = \text{Tr}(U\rho U^\dagger \mathcal{O})$

In this appendix, the Hessian operator of  $J$  is described and its eigenvalues and eigenvectors obtained. First, note that, since  $U^\dagger U = \mathbb{I}$ ,  $\delta U^\dagger U + U^\dagger \delta U = 0$  for any  $\delta U \in T_U \mathfrak{U}(N)$ , so  $\delta U^\dagger = -U^\dagger \delta U U^\dagger$ . Let  $J : \mathfrak{U}(N) \rightarrow \mathbb{R}$  be given by  $J(U) = \text{Tr}(U\rho U^\dagger \mathcal{O})$ . Then

$$d_U J(\delta U) = \Re \text{Tr}(\delta U \rho U^\dagger \mathcal{O} + U \rho \delta U^\dagger \mathcal{O}) = \Re \text{Tr}(\delta U \rho U^\dagger \mathcal{O} - U \rho U^\dagger \delta U U^\dagger \mathcal{O}) \quad (46a)$$

$$= \Re \text{Tr}([\rho, U^\dagger \mathcal{O} U] U^\dagger \delta U) = \langle U[U^\dagger \mathcal{O} U, \rho], \delta U \rangle \quad (46b)$$

so that  $\text{grad } J(U) = U[U^\dagger \mathcal{O} U, \rho] \in T_U \mathfrak{U}(N)$  and  $\text{grad } J$  is a vector field over  $\mathfrak{U}(N)$ , i.e.  $\text{grad } J \in \mathfrak{X}(\mathfrak{U}(N))$  [the  $C^\infty$  module of smooth vector fields on  $\mathfrak{U}(N)$ ]. For any  $X \in \mathfrak{X}(\mathfrak{U}(N))$ , the Hessian of  $J$  is given by  $\text{Hess}(X) = \nabla_X \text{grad } J$ . Because  $\mathfrak{U}(N)$  is endowed with the Riemannian metric induced from the real Hilbert-Schmidt inner product on  $\mathbb{C}^{N \times N}$ ,  $\nabla_X \text{grad } J = (d(\text{grad } J)(\bar{X}))^\mathcal{T}$  [17], where  $\bar{X}$  is an extension of  $X$  to  $\mathbb{C}^{N \times N}$  and  $\mathcal{T}$  denotes the tangential part. Now,

$$d_U \text{grad } J(\delta U) = \delta U[U^\dagger \mathcal{O} U, \rho] + U[\delta U^\dagger \mathcal{O} U, \rho] + U[U^\dagger \mathcal{O} \delta U, \rho] \quad (47)$$

and therefore,

$$\text{Hess}(X) = \nabla_X \text{grad } J = \frac{1}{2}U \left( U^\dagger X [U^\dagger \mathcal{O}U, \rho] + [U^\dagger \mathcal{O}U, \rho] X^\dagger U + 2[X^\dagger \mathcal{O}U, \rho] + 2[U^\dagger \mathcal{O}X, \rho] \right) \quad (48a)$$

$$= \frac{1}{2}U \left( U^\dagger X [U^\dagger \mathcal{O}U, \rho] - [U^\dagger \mathcal{O}U, \rho] U^\dagger X + 2[X^\dagger U U^\dagger \mathcal{O}U, \rho] + 2[U^\dagger \mathcal{O}U U^\dagger X, \rho] \right) \quad (48b)$$

$$= \frac{1}{2}U \left( U^\dagger X [U^\dagger \mathcal{O}U, \rho] - [U^\dagger \mathcal{O}U, \rho] U^\dagger X - 2[U^\dagger X U^\dagger \mathcal{O}U, \rho] + 2[U^\dagger \mathcal{O}U U^\dagger X, \rho] \right) \quad (48c)$$

$$= U \left( \frac{1}{2}[U^\dagger X, [U^\dagger \mathcal{O}U, \rho]] + [[U^\dagger \mathcal{O}U, U^\dagger X], \rho] \right). \quad (48d)$$

Any Hessian may be written as  $\langle Y, \text{Hess}(X) \rangle = XYJ - (\nabla_X Y)J$ . Note that  $XYJ$  is second order in  $J$  and first order in the manifold (i.e., relates to tangent spaces), while  $(\nabla_X Y)J$  is first order in  $J$  and second order in the manifold (i.e., relates to curvature of the manifold), so this second term vanishes at critical points and on flat spaces (when  $X$  and  $Y$  are constant vector fields). In the above case, for left invariant  $X$  and  $Y$ , it can be shown that  $XYJ = \langle Y, U[[U^\dagger \mathcal{O}U, U^\dagger X], \rho] \rangle$  and  $(\nabla_X Y)J = -\langle Y, \frac{1}{2}U[U^\dagger X, [U^\dagger \mathcal{O}U, \rho]] \rangle$ .

Suppose that  $U$  is a critical point of  $J$ , so that  $\text{grad } J(U) = U[U^\dagger \mathcal{O}U, \rho] = 0$  and  $\text{Hess}(X) = U[[U^\dagger \mathcal{O}U, U^\dagger X], \rho]$ . Let  $\Omega$  diagonalize  $\rho$  and  $\Gamma$  diagonalize  $\mathcal{O}$  such that the diagonal elements of  $\Lambda = \Omega \rho \Omega^\dagger$  and  $\Sigma = \Gamma \mathcal{O} \Gamma^\dagger$  are decreasing. Then  $U = \Gamma^\dagger V \mathcal{P} W^\dagger \Omega$  for some  $(V, W) \in \text{U}(\mathbf{m}) \oplus \text{U}(\mathbf{n})$  and some  $\mathcal{P} \in S_N$  [7]. We can then write

$$\text{Hess}(UY) = U[[\Omega^\dagger W \mathcal{P}^\dagger \Sigma \mathcal{P} W \Omega, Y], \Omega^\dagger \Lambda \Omega] = U \Omega^\dagger W [[\mathcal{P}^\dagger \Sigma \mathcal{P}, W^\dagger \Omega Y \Omega^\dagger W], \Lambda] \Omega W^\dagger. \quad (49)$$

Then  $\text{Hess}(UY) = \beta UY$  if and only if

$$\beta \tilde{Y} = [[\mathcal{P}^\dagger \Sigma \mathcal{P}, \tilde{Y}], \Lambda] \quad (50a)$$

$$\beta \tilde{Y}_{jk} = -(\lambda_j - \lambda_k)(\sigma_{\mathcal{P}(j)} - \sigma_{\mathcal{P}(k)}) \tilde{Y}_{jk} \quad (50b)$$

where  $\tilde{Y} := W^\dagger \Omega Y \Omega^\dagger W$ . So either  $\beta = -(\lambda_j - \lambda_k)(\sigma_{\mathcal{P}(j)} - \sigma_{\mathcal{P}(k)})$  or  $\tilde{Y}_{jk} = 0$ . Then an orthonormal set of solutions is given by

$$\beta = 0 \quad \tilde{Y} = i|l\rangle\langle l| \quad (51a)$$

$$\beta = -(\lambda_j - \lambda_k)(\sigma_{\mathcal{P}(j)} - \sigma_{\mathcal{P}(k)}) \quad \tilde{Y} = \frac{i}{\sqrt{2}}(|j\rangle\langle k| + |k\rangle\langle j|) \quad (51b)$$

$$\beta = -(\lambda_j - \lambda_k)(\sigma_{\mathcal{P}(j)} - \sigma_{\mathcal{P}(k)}) \quad \tilde{Y} = \frac{1}{\sqrt{2}}(|j\rangle\langle k| - |k\rangle\langle j|) \quad (51c)$$

for  $1 \leq l \leq N$  and  $1 \leq j < k \leq N$ . These  $\beta$  and the corresponding vectors  $UY = U \Omega^\dagger W \tilde{Y} W^\dagger \Omega$  are the eigenvalues and eigenvectors of Hess at the critical point  $U$ .

## APPENDIX C. THE SHAPE OPERATORS OF THE CRITICAL SUBMANIFOLDS

Consider a critical submanifold for some permutation matrix  $\mathcal{P}$  and degeneracy structures  $\mathbf{m}$  and  $\mathbf{n}$  of  $\mathcal{O}$  and  $\rho$ , respectively. Then, let  $\Omega \in \mathrm{U}(N)$  and  $\Gamma \in \mathrm{U}(N)$  diagonalize  $\rho$  and  $\mathcal{O}$ , so that  $\Omega\rho\Omega^\dagger = \Lambda$  and  $\Gamma\mathcal{O}\Gamma^\dagger$  with  $\lambda_1 \geq \dots \geq \lambda_N$  and  $\sigma_1 \geq \dots \geq \sigma_N$ . The critical submanifold is then  $C = \Gamma^\dagger \mathrm{U}(\mathbf{n})\mathcal{P}\mathrm{U}(\mathbf{m})\Omega \subset \mathrm{U}(N)$ . At any  $U = \Gamma^\dagger V\mathcal{P}W^\dagger\Omega$  in this critical submanifold, the Hessian eigenvectors identified in (51) corresponding to zero eigenvalue describe an orthonormal basis for the tangent space of  $C$ ,  $\mathrm{T}_U C$ , and the eigenvectors corresponding to nonzero eigenvalues form an orthonormal basis for the normal space of  $C$  at  $U$ ,  $(\mathrm{T}_U C)^\perp$ . While such a critical submanifold  $C$  is not totally geodesic in general, it is the case that the geodesics along these particular basis directions for  $\mathrm{T}_U C$  remain on  $C$ . To see this, first note that each of these basis vectors  $X$  has the property that either  $[U^\dagger X, U^\dagger \mathcal{O}U] = 0$  or  $[U^\dagger X, \rho] = 0$ . This follows from the fact that, when  $\tilde{X} = W^\dagger \Omega U^\dagger X \Omega^\dagger W \in \mathfrak{u}(N)$ , either  $\tilde{X} = i|l\rangle\langle l|$  for some  $l = 1, \dots, N$  so that  $X$  commutes with both  $U^\dagger \mathcal{O}U$  and  $\rho$ , or  $\tilde{X} = z|j\rangle\langle k| - \bar{z}|k\rangle\langle j|$  for some  $|z|^2 = 1/2$  and  $j < k$  with either  $\lambda_j = \lambda_k$  or  $\sigma_{\mathcal{P}(j)} = \sigma_{\mathcal{P}(k)}$ . When  $U^\dagger X$  commutes with  $U^\dagger \mathcal{O}U$ , then  $\mathcal{P}\tilde{X}\mathcal{P}^\dagger$  commutes with  $\Sigma$ , so that  $\mathcal{P}\tilde{X}\mathcal{P}^\dagger \in \mathfrak{u}(\mathbf{m})$ . Then the curve  $\gamma(s) = e^{sX}U^\dagger = \Gamma^\dagger V e^{s\mathcal{P}\tilde{X}\mathcal{P}^\dagger} \mathcal{P}W^\dagger\Omega$  is a geodesic in  $C$  in the direction  $\gamma'(0) = X$ , which can be thought of in terms of the path through  $\mathrm{U}(\mathbf{m}) \oplus \mathrm{U}(\mathbf{n})$ ,  $(V(s), W(s)) = (V e^{s\mathcal{P}\tilde{X}\mathcal{P}^\dagger}, W)$ . Likewise, when  $U^\dagger X$  commutes with  $\rho$ , then  $\tilde{X}$  commutes with  $\Lambda$ , so that  $\tilde{X} \in \mathfrak{u}(\mathbf{n})$ . Then the curve  $\gamma(s) = U e^{sU^\dagger X} = \Gamma^\dagger V \mathcal{P} e^{s\tilde{X}} W^\dagger\Omega$  is a geodesic in  $C$  in the direction  $\gamma'(0) = X$ , which can be thought of in terms of the path through  $\mathrm{U}(\mathbf{m}) \oplus \mathrm{U}(\mathbf{n})$ ,  $(V(s), W(s)) = (V, W e^{-s\tilde{X}})$ . Note that, while the geodesic curve is uniquely defined for a given basis vector  $X$ , the corresponding path through  $\mathrm{U}(\mathbf{m}) \oplus \mathrm{U}(\mathbf{n})$  is not unique because of the ambiguity in associating  $(V, W) \in \mathrm{U}(\mathbf{m}) \oplus \mathrm{U}(\mathbf{n})$  to a given  $U \in C$  afforded by the stabilizer subgroup  $\mathrm{Stab}_{\mathrm{U}(\mathbf{m}) \oplus \mathrm{U}(\mathbf{n})}(\mathcal{P}) \cong \mathrm{U}(\mathbf{K})$ .

For any two of these basis vectors for  $\mathrm{T}_U C$ , say  $X$  and  $Y$ , the second fundamental form  $S(X, Y)$  at  $U$  is defined to be  $(\nabla_{\hat{X}} \hat{Y})_U^\perp$ , where  $\hat{X}$  and  $\hat{Y}$  are smooth local extensions of  $X$  and  $Y$ . This can also be calculated by extending  $Y$  along the geodesic in the direction of  $X$  and taking the normal part of the covariant derivative of this field. Letting  $\tilde{Y} = W^\dagger \Omega U^\dagger Y \Omega^\dagger W \in \mathfrak{u}(N)$ , when  $U^\dagger X$  commutes with  $U^\dagger \mathcal{O}U$ , the vector  $Y$  can be extended along the geodesic  $\gamma(s) = e^{sX}U^\dagger$  by  $\hat{Y}(s) = \gamma(s)\Omega^\dagger W \tilde{Y} W^\dagger\Omega = \Gamma^\dagger V e^{s\mathcal{P}\tilde{X}\mathcal{P}^\dagger} \mathcal{P}\tilde{Y} W^\dagger\Omega = \Gamma^\dagger V \mathcal{P} e^{s\tilde{X}} \tilde{Y} W^\dagger\Omega$  where the path through  $\mathrm{U}(\mathbf{m}) \oplus \mathrm{U}(\mathbf{n})$   $(V(s), W(s)) = (V e^{s\mathcal{P}\tilde{X}\mathcal{P}^\dagger}, W)$  is used to define the extension. Likewise, when  $U^\dagger X$  commutes with  $\rho$ , the vector  $Y$  can be extended along the geodesic  $\gamma(s) = U e^{sU^\dagger X}$  by  $\hat{Y}(s) = \Gamma^\dagger V \mathcal{P} \tilde{Y} e^{s\tilde{X}} W^\dagger\Omega$  by using  $(V(s), W(s)) = (V, W e^{-s\tilde{X}})$ . Since the paths through  $\mathrm{U}(\mathbf{m}) \oplus \mathrm{U}(\mathbf{n})$  are not uniquely defined, these extensions of  $Y$  along  $\gamma$  are not unique. Indeed, this is clear in the cases where  $U^\dagger X$  commutes with both  $U^\dagger \mathcal{O}U$  and  $\rho$ , as the two extensions of  $Y$  described here are different for any  $Y$  such that  $[\tilde{Y}, \tilde{X}] \neq 0$ . However, since the second fundamental form is tensorial [20], the choice of extension of  $Y$  will ultimately not be important.

The covariant derivative of  $\hat{Y}(s)$  can be computed simply as the ordinary derivative of  $\hat{Y}$  (as a path through  $\mathbb{C}^{N \times N}$ ) followed by projection down to the tangent space  $\mathrm{T}_{\gamma(s)}\mathrm{U}(N)$ . Then, using  $^\top$  to denote the tangential component, the covariant derivatives for the particular extensions  $\hat{Y}(s)$

defined above are

$$\nabla_{\hat{X}} \hat{Y} = \left( \frac{d}{ds} \hat{Y}(s) \Big|_{s=0} \right)^\top = \begin{cases} \left( \Gamma^\dagger V \mathcal{P} \tilde{X} \tilde{Y} W^\dagger \Omega \right)^\top & \text{for } [U^\dagger X, U^\dagger \mathcal{O}U] = 0 \\ \left( \Gamma^\dagger V \mathcal{P} \tilde{Y} \tilde{X} W^\dagger \Omega \right)^\top & \text{for } [U^\dagger X, \rho] = 0 \end{cases} \quad (52a)$$

$$= \begin{cases} \frac{1}{2} \left( \Gamma^\dagger V \mathcal{P} \tilde{X} \tilde{Y} W^\dagger \Omega - U \Omega^\dagger W^\dagger \tilde{Y} \tilde{X} \mathcal{P}^\dagger V^\dagger \Gamma U \right) & \text{for } [U^\dagger X, U^\dagger \mathcal{O}U] = 0 \\ \frac{1}{2} \left( \Gamma^\dagger V \mathcal{P} \tilde{Y} \tilde{X} W^\dagger \Omega - U \Omega^\dagger W^\dagger \tilde{X} \tilde{Y} \mathcal{P}^\dagger V^\dagger \Gamma U \right) & \text{for } [U^\dagger X, \rho] = 0 \end{cases} \quad (52b)$$

$$= \begin{cases} \frac{1}{2} \Gamma^\dagger V \mathcal{P} [\tilde{X}, \tilde{Y}] W^\dagger \Omega & \text{for } [U^\dagger X, U^\dagger \mathcal{O}U] = 0 \\ -\frac{1}{2} \Gamma^\dagger V \mathcal{P} [\tilde{X}, \tilde{Y}] W^\dagger \Omega & \text{for } [U^\dagger X, \rho] = 0. \end{cases} \quad (52c)$$

Then  $S(X, Y) = (\nabla_{\hat{X}} \hat{Y})^\perp$  is the normal component, where the normal space  $(T_U C)^\perp$  is spanned by the vectors  $Z = \Gamma^\dagger V \mathcal{P} \tilde{Z} W^\dagger \Omega$  where  $\tilde{Z} = z|j\rangle\langle k| - \bar{z}|k\rangle\langle j|$  such that  $|z|^2 = 1/2$  and  $j$  and  $k$  are such that  $\lambda_j \neq \lambda_k$  and  $\sigma_{\mathcal{P}(j)} \neq \sigma_{\mathcal{P}(k)}$ . Now, observe that if  $\tilde{X} = i|l\rangle\langle l|$  and  $\tilde{Y} = i|m\rangle\langle m|$ , then  $[\tilde{X}, \tilde{Y}] = 0$ . And if  $\tilde{X} = i|l\rangle\langle l|$  and  $\tilde{Y} = w|j\rangle\langle k| - \bar{w}|k\rangle\langle j|$  for  $j < k$ , then  $[\tilde{X}, \tilde{Y}] = (\delta_{jl} - \delta_{kl})(i\bar{w})|j\rangle\langle k| - \overline{(i\bar{w})}|k\rangle\langle j|$  so that  $\nabla_{\hat{X}} \hat{Y} \in T_U C$ . So whenever  $X$  is such that  $\tilde{X} = i|l\rangle\langle l|$ ,  $S(X, Y) = 0$  for all  $Y \in T_U C$ . Moreover, if  $\tilde{X} = v|l\rangle\langle m| - \bar{v}|m\rangle\langle l|$  and  $\tilde{Y} = w|j\rangle\langle k| - \bar{w}|k\rangle\langle j|$ , then

$$\begin{aligned} [\tilde{X}, \tilde{Y}] &= \delta_{jm}(vw|l\rangle\langle k| - \bar{v}\bar{w}|k\rangle\langle l|) + \delta_{km}(\bar{v}w|j\rangle\langle l| - v\bar{w}|l\rangle\langle j|) \\ &\quad + \delta_{kl}(\bar{v}\bar{w}|m\rangle\langle j| - vw|j\rangle\langle m|) + \delta_{jl}(v\bar{w}|k\rangle\langle m| - \bar{v}w|m\rangle\langle k|). \end{aligned} \quad (53)$$

Consequently, if  $U^\dagger X$  and  $U^\dagger Y$  both commute with  $U^\dagger \mathcal{O}U$ , then  $\nabla_{\hat{X}} \hat{Y}$  from (52c) also commutes with  $U^\dagger \mathcal{O}U$  and therefore lies in  $T_U C$ , so that  $S(X, Y) = (\nabla_{\hat{X}} \hat{Y})^\perp = 0$ . Likewise, if  $U^\dagger X$  and  $U^\dagger Y$  both commute with  $\rho$ , then  $\nabla_{\hat{X}} \hat{Y}$  commutes with  $\rho$  so that again  $S(X, Y) = (\nabla_{\hat{X}} \hat{Y})^\perp = 0$ . So among these basis vectors for  $T_U C$ , the only pairs  $X$  and  $Y$  for which  $S(X, Y)$  can be non-zero are those in which one of the pair  $\{U^\dagger X, U^\dagger Y\}$  commutes with  $U^\dagger \mathcal{O}U$ , but not with  $\rho$ , and the other commutes with  $\rho$ , but not with  $U^\dagger \mathcal{O}U$ . Then the basis vectors for  $T_U C$  can be divided into three categories: (1) those for which  $U^\dagger X$  commutes with  $U^\dagger \mathcal{O}U$  but not  $\rho$ , of which there are  $\sum m_i^2 - \sum k_{ij}^2$ ; (2) those for which  $U^\dagger X$  commutes with  $\rho$  but not  $U^\dagger \mathcal{O}U$ , of which there are  $\sum n_j^2 - \sum k_{ij}^2$ ; and (3) the remaining  $\sum k_{ij}^2$  basis vectors for which  $U^\dagger X$  commutes with both  $U^\dagger \mathcal{O}U$  and  $\rho$ .

Taking any normal vector  $Z \in (T_U C)^\perp$ , the *shape operator* (or Weingarten map) [20]  $A_Z : T_U C \rightarrow T_U C$ , defined by  $\langle A_Z X, Y \rangle = \langle S(X, Y), Z \rangle$ , can be represented in block form using these three categories of tangent vectors as

$$A_Z = \begin{bmatrix} 0 & B & 0 \\ B^T & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (54)$$

where  $B$  is an  $(\sum m_i^2 - \sum k_{ij}^2) \times (\sum n_j^2 - \sum k_{ij}^2)$  matrix. The eigenvalues of  $A_Z$  are the *principal curvatures* of the submanifold  $C$  at the point  $U$  with respect to the normal vector  $Z$ . Because of the indicated block structure, the nonzero eigenvalues of  $A_Z$  come in positive-negative pairs  $\pm\eta$ , where

$\eta$  is a nonzero eigenvalue of  $B^T B$ . Moreover, these eigenvalues can be bounded by observing that the squared Hilbert-Schmidt norm  $\|B\|^2 = \sum_{j < m} \frac{|\tilde{Z}_{jm}|^2}{2} (k_{q_m t_j} + k_{p_j u_m})$  where  $\lambda_m = \tilde{\lambda}_{q_m}$ ,  $\lambda_j = \tilde{\lambda}_{p_j}$ ,  $\sigma_{\mathcal{P}(j)} = \tilde{\sigma}_{t_j}$ ,  $\sigma_{\mathcal{P}(m)} = \tilde{\sigma}_{u_m}$ , and  $\tilde{\lambda}_1 > \dots > \tilde{\lambda}_s$  and  $\tilde{\sigma}_1 > \dots > \tilde{\sigma}_r$  are the distinct eigenvalues of  $\rho$  and  $\mathcal{O}$  with multiplicities  $m_1, \dots, m_s$  and  $n_1, \dots, n_r$ . Most importantly,  $\text{Tr}(A_Z) = 0$  for every normal vector  $Z$ , so that the mean curvature vector field [16]  $H = \sum S(X_i, X_i)$  for any orthonormal basis  $\{X_i\}$  of  $T_U C$ , is zero. Therefore every critical submanifold of  $J(U) = \text{Tr}(U \rho U^\dagger \mathcal{O})$  is a *minimal* submanifold of  $U(N)$  [20].

**C.1. Examples.** We now return to the examples considered in Sections 3.1 and 4.3 and compute the shape operators of the critical submanifolds.

**Example 7** (Maximum Submanifold of  $P_{i \rightarrow f}$ ). *From the contingency table for  $P_{i \rightarrow f} = 1$  (see Section 3.1), it is clear that  $\sum m_i^2 - \sum k_{ij}^2 = \sum n_j^2 - \sum k_{ij}^2 = 0$ , so that the matrix  $B$  in the shape operator  $A_Z$  is  $0 \times 0$ . Therefore  $A_Z$  is always the zero operator for all normal vectors  $Z$ , meaning that this critical submanifold is totally geodesic [20].*

**Example 8** (Minimum Submanifold of  $P_{i \rightarrow f}$ ). *Consider the case where  $\sigma_{\mathcal{P}(2)} = 1$  and  $\sigma_{\mathcal{P}(j)} = 0$  for all  $j \neq 2$ . The first category basis vectors are then  $X_j = \Gamma^\dagger V \mathcal{P} \tilde{X}_j W^\dagger \Omega$ , where  $\tilde{X}_j = w_j |1\rangle\langle j| - \bar{w}_j |j\rangle\langle 1|$  for  $j = 3, \dots, N$  and for  $w_j = \frac{1}{\sqrt{2}}$  and  $\bar{w}_j = \frac{i}{\sqrt{2}}$ . Likewise, the second category basis vectors are  $Y_j = \Gamma^\dagger V \mathcal{P} \tilde{Y}_j W^\dagger \Omega$  for  $\tilde{Y}_j = v_j |2\rangle\langle j| - \bar{v}_j |j\rangle\langle 2|$  for  $j = 3, \dots, N$ , and for  $v_j = \frac{1}{\sqrt{2}}$  and  $\bar{v}_j = \frac{i}{\sqrt{2}}$ . So  $\nabla_{\tilde{X}_j} \hat{Y}_k = \delta_{jk} \frac{1}{2} \Gamma^\dagger V \mathcal{P} [\tilde{X}_j, \tilde{Y}_j] W^\dagger \Omega$ , where  $[\tilde{X}_j, \tilde{Y}_j] = -w_j \bar{v}_j |1\rangle\langle 2| + \bar{w}_j v_j |2\rangle\langle 1|$ . Then for any normal vector  $Z = \Gamma^\dagger V \mathcal{P} \tilde{Z} W^\dagger \Omega$  with  $\tilde{Z} = z |1\rangle\langle 2| - \bar{z} |2\rangle\langle 1|$ ,  $\langle \nabla_{\tilde{X}} \hat{Y}, Z \rangle = \frac{1}{2} \langle [\tilde{X}, \tilde{Y}], \tilde{Z} \rangle = -\Re(\bar{w}_j v_j z)$ , so the  $(2N - 4) \times (2N - 4)$  matrix  $B$  in the shape operator  $A_Z$  is*

$$B = -\frac{1}{2} \mathbb{I}_{N-2} \otimes \begin{bmatrix} \Re(z) & \Im(z) \\ -\Im(z) & \Re(z) \end{bmatrix}. \quad (55)$$

*Then  $B^T B = \frac{|z|^2}{4} \mathbb{I}_{2N-4}$ , so that  $A_Z$  has eigenvalue  $\eta = \frac{|z|}{2}$  with multiplicity  $2N - 4$ ,  $\eta = -\frac{|z|}{2}$  with multiplicity  $2N - 4$ , and  $\eta = 0$  with multiplicity  $N^2 - 4N + 6$ . When  $Z$  is normalized with respect to the Hilbert-Schmidt norm,  $|z|^2 = 1/2$ , so the nonzero eigenvalues of  $A_Z$  are  $\eta = \pm \frac{1}{2\sqrt{2}}$ .*

**Example 9** (Fully Non-degenerate  $\rho$  and  $\mathcal{O}$ ). *In the case where  $\rho$  and  $\mathcal{O}$  are both fully non-degenerate,  $m_1 = \dots = m_N = 1$ ,  $n_1 = \dots = n_N = 1$  and  $\mathbf{K} = \mathcal{P}^\dagger$ . Then for any critical submanifold  $\text{Orb}(\mathcal{P})$ ,  $\sum m_i^2 - \sum k_{ij}^2 = \sum n_j^2 - \sum k_{ij}^2 = 0$ , so that the matrix  $B$  in the shape operator  $A_Z$  is  $0 \times 0$ . Therefore  $A_Z$  is always the zero operator for all normal vectors  $Z$ , so that every critical submanifold of  $J$  is totally geodesic when  $\rho$  and  $\mathcal{O}$  are fully non-degenerate.*

## APPENDIX D. IN SUPPORT OF CONJECTURE 1

In this appendix we build the case for Conjecture 1. Let  $U = \Gamma^\dagger V \mathcal{P} W^\dagger \Omega$  be a critical point, where  $(V, W) \in \mathbf{U}(\mathbf{m}) \oplus \mathbf{U}(\mathbf{n})$  and  $\Omega, \Gamma \in \mathbf{U}(N)$  diagonalize  $\rho$  and  $\mathcal{O}$ , respectively, i.e.  $\Omega \rho \Omega^\dagger = \Lambda$  and  $\Gamma \mathcal{O} \Gamma^\dagger = \Sigma$ . Also let  $\tilde{A} = \sum_{l=1}^L \alpha_l \tilde{A}_l$ , where  $\sum \alpha_l^2 = 1$ ,  $\tilde{A}_l = z_l |j_l\rangle \langle k_l| - \bar{z}_l |k_l\rangle \langle j_l|$ ,  $|z_l|^2 = 1/2$ , and  $\{j_1, k_2, \dots, j_L, k_L\}$  is a set of  $2L$  distinct indices in  $\{1, \dots, N\}$  such that the Hessian eigenvalue  $\beta_{j_l k_l} = -(\lambda_{j_l} - \lambda_{k_l})(\sigma_{\mathcal{P}(j_l)} - \sigma_{\mathcal{P}(k_l)})$  is nonzero. As shown in Appendix B, for the Hessian at a critical point, every eigenspace corresponding to a nonzero eigenvalue is spanned by eigenvectors of the form  $z|j\rangle \langle k| - \bar{z}|k\rangle \langle j|$  with  $j < k$ . Then  $\exp(sA)\rho \exp(-sA) = \Omega^\dagger W \exp(s\tilde{A})\Lambda \exp(-s\tilde{A})W^\dagger \Omega$ . Now, defining  $\tilde{A}_l^+ = z_l |j_l\rangle \langle k_l| + \bar{z}_l |k_l\rangle \langle j_l|$  and  $\tilde{A}^+ = \sum \alpha_l \tilde{A}_l^+$ , it is easily verified that

$$[\tilde{A}, \Lambda] = \sum_{l=1}^L (\lambda_{k_l} - \lambda_{j_l}) \alpha_l \tilde{A}_l^+ \quad (56a)$$

$$[\tilde{A}, \tilde{A}_l^+] = \alpha_l (|j_l\rangle \langle j_l| - |k_l\rangle \langle k_l|) \quad (56b)$$

$$[\tilde{A}, |j_l\rangle \langle j_l| - |k_l\rangle \langle k_l|] = -2\alpha_l \tilde{A}_l^+, \quad (56c)$$

so that  $[\tilde{A}, [\tilde{A}, \tilde{A}_l^+]] = -2\alpha_l^2 \tilde{A}_l^+$ , and  $[\tilde{A}, [\tilde{A}, |j_l\rangle \langle j_l| - |k_l\rangle \langle k_l|]] = -2\alpha_l^2 (|j_l\rangle \langle j_l| - |k_l\rangle \langle k_l|)$ . Then, using the notation  $\text{ad}_X$  for the adjoint operator  $\text{ad}_X(Y) = [X, Y]$  [19],

$$\exp(s\tilde{A})\Lambda \exp(-s\tilde{A}) = \exp(s \text{ad}_{\tilde{A}})\Lambda = \sum_{m=0}^{\infty} \frac{s^m \text{ad}_{\tilde{A}}^m(\Lambda)}{m!} \quad (57a)$$

$$= \Lambda + \sum_{m=1}^{\infty} \frac{s^{2m} \text{ad}_{\tilde{A}}^{2m}(\Lambda)}{(2m)!} + \sum_{m=0}^{\infty} \frac{s^{2m+1} \text{ad}_{\tilde{A}}^{2m+1}(\Lambda)}{(2m+1)!} \quad (57b)$$

$$= \Lambda + \sum_{l=1}^L \alpha_l^2 (\lambda_{k_l} - \lambda_{j_l}) \sum_{m=1}^{\infty} \frac{s^{2m} \text{ad}_{\tilde{A}}^{2m-2} (|j_l\rangle \langle j_l| - |k_l\rangle \langle k_l|)}{(2m)!} \\ + \sum_{l=1}^L \alpha_l (\lambda_{k_l} - \lambda_{j_l}) \sum_{m=0}^{\infty} \frac{s^{2m+1} \text{ad}_{\tilde{A}}^{2m}(\tilde{A}_l^+)}{(2m+1)!} \quad (57c)$$

$$= \Lambda + \sum_{l=1}^L (\lambda_{k_l} - \lambda_{j_l}) \sum_{m=1}^{\infty} \frac{(\alpha_l s)^{2m} (-2)^{m-1}}{(2m)!} (|j_l\rangle \langle j_l| - |k_l\rangle \langle k_l|) \\ + \sum_{l=1}^L (\lambda_{k_l} - \lambda_{j_l}) \sum_{m=0}^{\infty} \frac{(\alpha_l s)^{2m+1} (-2)^m}{(2m+1)!} \tilde{A}_l^+ \quad (57d)$$

$$= \Lambda + \sum_l \frac{\lambda_{k_l} - \lambda_{j_l}}{2} (1 - \cos(\sqrt{2}\alpha_l s)) (|j_l\rangle \langle j_l| - |k_l\rangle \langle k_l|) \\ + \sum_l \frac{\lambda_{k_l} - \lambda_{j_l}}{\sqrt{2}} \sin(\sqrt{2}\alpha_l s) \tilde{A}_l^+ \quad (57e)$$

so that

$$[U^\dagger \mathcal{O}U, e^{sA} \rho e^{-sA}] = \Omega^\dagger W[\mathcal{P}^\dagger \Sigma \mathcal{P}, \exp(s\tilde{A}) \Lambda \exp(-s\tilde{A})] W^\dagger \Omega \quad (58a)$$

$$\begin{aligned} &= \Omega^\dagger W \left( [\mathcal{P}^\dagger \Sigma \mathcal{P}, \Lambda] + \sum_l \frac{\lambda_{k_l} - \lambda_{j_l}}{2} (1 - \cos(\sqrt{2}\alpha_l s)) [\mathcal{P}^\dagger \Sigma \mathcal{P}, |j_l\rangle\langle j_l| - |k_l\rangle\langle k_l|] \right. \\ &\quad \left. + \sum_l \frac{\lambda_{k_l} - \lambda_{j_l}}{\sqrt{2}} \sin(\sqrt{2}\alpha_l s) [\mathcal{P}^\dagger \Sigma \mathcal{P}, \tilde{A}_l^+] \right) W^\dagger \Omega \end{aligned} \quad (58b)$$

$$= \sum_l \frac{(\lambda_{k_l} - \lambda_{j_l})(\sigma_{\mathcal{P}(k_l)} - \sigma_{\mathcal{P}(j_l)})}{\sqrt{2}} \sin(\sqrt{2}\alpha_l s) A_l \quad (58c)$$

$$= - \sum_l \frac{\beta_{j_l k_l}}{\sqrt{2}} \sin(\sqrt{2}\alpha_l s) A_l \quad (58d)$$

where  $\beta_{jk}$  is the Hessian eigenvalue corresponding to the eigenvector  $U(z|j\rangle\langle k| - \bar{z}|k\rangle\langle j|)$ . It follows that, for this choice of  $A$ ,  $f(s) = \sum_l \beta_{j_l k_l}^2 \sin^2(\sqrt{2}\alpha_l s)/2$  for all  $s$ . Now, let  $q(s) = \sin^2(\sqrt{2}\alpha_l s) - \alpha_l^2 \sin^2(\sqrt{2}s)$ . Then  $q'(s) = \sqrt{2}\alpha_l \sin(2\sqrt{2}\alpha_l s) - \sqrt{2}\alpha_l^2 \sin(2\sqrt{2}s)$ , and  $q''(s) = 4\alpha_l^2 [\cos(2\sqrt{2}\alpha_l s) - \cos(2\sqrt{2}s)]$ . Since  $q'(0) = 0$  and  $q''(s) \geq 0$  for all  $s \in [0, \pi/(2\sqrt{2})]$ ,  $q'(s) \geq 0$  for all  $s \in [0, \pi/(2\sqrt{2})]$ . And since  $q(0) = 0$ , this implies that  $q(s) \geq 0$  for all  $s \in [0, \pi/(2\sqrt{2})]$ . Therefore,  $f(s) \geq \sum_l \beta_{j_l k_l}^2 \alpha_l^2 \sin^2(\sqrt{2}s)/2 \geq \beta_{\min}^2 \sin^2(\sqrt{2}s)/2$ , since  $\sum \alpha_l^2 = 1$ . So for all unit normal vectors  $UA$  of this form,  $f(s)$  satisfies the conjecture. Moreover, when  $\tilde{A} = z|j\rangle\langle k| - \bar{z}|k\rangle\langle j|$  for some  $j \neq k$ ,  $f(s) = \beta_{jk}^2 \sin^2(\sqrt{2}s)/2$ , and in the particular case when  $j$  and  $k$  are such that  $UA$  is an eigenvector corresponding to the smallest eigenvalue (in absolute value), then  $f(s) = \beta_{\min}^2 \sin^2(\sqrt{2}s)/2$  for all  $s$ , so this conjectured lower bound is attained.

Now, for that same critical point  $U = \Gamma^\dagger V \mathcal{P} W^\dagger \Omega$ , let  $\mathcal{N}(\text{Hess})^\perp \subset T_U U(N)$  be the orthogonal complement of the null space of the Hessian at  $U$ , i.e. the span of the Hessian eigenvectors corresponding to non-zero eigenvalues. Then let  $\mathcal{S} \subset U^\dagger Z \subset \mathfrak{u}(N)$  be the unit sphere within  $U^\dagger Z$ . For any  $s \in [0, \pi/(2\sqrt{2})]$ , let  $\zeta_s : \mathcal{S} \rightarrow \mathbb{R}$  be given by  $\zeta_s(A) = \|\text{grad } J(U \exp(sA))\|^2$ . Then, using the integral expression for the derivative of the matrix exponential [21, App. B], the differential of  $\zeta_s$  is found to be

$$d_A \zeta_s(\delta A) = 2 \left\langle \left[ e^{-sA} U^\dagger \mathcal{O}U e^{sA}, \rho \right], \left[ \left[ e^{-sA} U^\dagger \mathcal{O}U e^{sA}, s \int_0^1 e^{-srA} \delta A e^{srA} dr \right], \rho \right] \right\rangle \quad (59a)$$

$$= 2s \left\langle \int_0^1 e^{srA} \left[ e^{-sA} U^\dagger \mathcal{O}U e^{sA}, \left[ \left[ e^{-sA} U^\dagger \mathcal{O}U e^{sA}, \rho \right], \rho \right] \right] e^{-srA} dr, \delta A \right\rangle \quad (59b)$$

$$= 2s \left\langle \int_0^1 e^{-s(1-r)\tilde{A}} \left[ \mathcal{P}^\dagger \Sigma \mathcal{P}, \left[ \left[ \mathcal{P}^\dagger \Sigma \mathcal{P}, e^{s\tilde{A}} \Lambda e^{-s\tilde{A}} \right], e^{s\tilde{A}} \Lambda e^{-s\tilde{A}} \right] \right] e^{s(1-r)\tilde{A}} dr, W^\dagger \Omega \delta A \Omega^\dagger W \right\rangle, \quad (59c)$$



where  $\tilde{A} = W^\dagger \Omega A \Omega^\dagger W$ . When  $\tilde{A} = \sum_l \alpha_l \tilde{A}_l$  with  $\tilde{A}_l = z_l |j_l\rangle \langle k_l| - \bar{z}_l |k_l\rangle \langle j_l|$  as above, we have seen that  $[\mathcal{P}^\dagger \Sigma \mathcal{P}, e^{s\tilde{A}} \Lambda e^{-s\tilde{A}}] = \sum_l \beta_{j_l k_l} \sin(\sqrt{2}\alpha_l s) \tilde{A}_l / \sqrt{2}$ . Then

$$\begin{aligned} & \left[ [\mathcal{P}^\dagger \Sigma \mathcal{P}, e^{s\tilde{A}} \Lambda e^{-s\tilde{A}}], e^{s\tilde{A}} \Lambda e^{-s\tilde{A}} \right] \\ &= \sum_{l=1}^L \frac{\beta_{j_l k_l}}{\sqrt{2}} \sin(\sqrt{2}\alpha_l s) e^{s\tilde{A}} [\tilde{A}, \Lambda] e^{-s\tilde{A}} \end{aligned} \quad (60a)$$

$$= \sum_{l=1}^L \alpha_l (\lambda_{k_l} - \lambda_{j_l}) \frac{\beta_{j_l k_l}}{\sqrt{2}} \sin(\sqrt{2}\alpha_l s) e^{s\tilde{A}} \tilde{A}_l^+ e^{-s\tilde{A}} \quad (60b)$$

$$= \sum_{l=1}^L \alpha_l (\lambda_{k_l} - \lambda_{j_l}) \frac{\beta_{j_l k_l}}{\sqrt{2}} \sin(\sqrt{2}\alpha_l s) \sum_{m=0}^{\infty} \frac{s^m \text{ad}_{\tilde{A}}^m(\tilde{A}_l^+)}{m!} \quad (60c)$$

$$\begin{aligned} &= \sum_{l=1}^L \alpha_l (\lambda_{k_l} - \lambda_{j_l}) \frac{\beta_{j_l k_l}}{\sqrt{2}} \sin(\sqrt{2}\alpha_l s) \sum_{m=0}^{\infty} \frac{s^{2m} \text{ad}_{\tilde{A}}^{2m}(\tilde{A}_l^+)}{(2m)!} \\ &\quad + \sum_{l=1}^L \alpha_l (\lambda_{k_l} - \lambda_{j_l}) \frac{\beta_{j_l k_l}}{\sqrt{2}} \sin(\sqrt{2}\alpha_l s) \sum_{m=0}^{\infty} \frac{s^{2m+1} \text{ad}_{\tilde{A}}^{2m+1}(\tilde{A}_l^+)}{(2m+1)!} \end{aligned} \quad (60d)$$

$$\begin{aligned} &= \sum_{l=1}^L \alpha_l (\lambda_{k_l} - \lambda_{j_l}) \frac{\beta_{j_l k_l}}{\sqrt{2}} \sin(\sqrt{2}\alpha_l s) \sum_{m=0}^{\infty} \frac{(\alpha_l s)^{2m} (-2)^m}{(2m)!} \tilde{A}_l^+ \\ &\quad + \sum_{l=1}^L \alpha_l (\lambda_{k_l} - \lambda_{j_l}) \frac{\beta_{j_l k_l}}{\sqrt{2}} \sin(\sqrt{2}\alpha_l s) \sum_{m=0}^{\infty} \frac{(\alpha_l s)^{2m+1} (-2)^m}{(2m+1)!} (|j_l\rangle \langle j_l| - |k_l\rangle \langle k_l|) \end{aligned} \quad (60e)$$

$$= \sum_{l=1}^L \alpha_l (\lambda_{k_l} - \lambda_{j_l}) \frac{\beta_{j_l k_l}}{2} \left\{ \frac{1}{\sqrt{2}} \sin(2\sqrt{2}\alpha_l s) \tilde{A}_l^+ + \sin^2(\sqrt{2}\alpha_l s) (|j_l\rangle \langle j_l| - |k_l\rangle \langle k_l|) \right\}, \quad (60f)$$

so that

$$\begin{aligned} & \int_0^1 e^{-s(1-r)\tilde{A}} \left[ \mathcal{P}^\dagger \Sigma \mathcal{P}, \left[ [\mathcal{P}^\dagger \Sigma \mathcal{P}, e^{s\tilde{A}} \Lambda e^{-s\tilde{A}}], e^{s\tilde{A}} \Lambda e^{-s\tilde{A}} \right] \right] e^{s(1-r)\tilde{A}} dr \\ &= - \sum_{l=1}^L \alpha_l \frac{\beta_{j_l k_l}^2}{2\sqrt{2}} \sin(2\sqrt{2}\alpha_l s) \int_0^1 e^{-s(1-r)\tilde{A}} \tilde{A}_l e^{s(1-r)\tilde{A}} dr \end{aligned} \quad (61a)$$

$$= - \sum_{l=1}^L \alpha_l \frac{\beta_{j_l k_l}^2}{2\sqrt{2}} \sin(2\sqrt{2}\alpha_l s) \tilde{A}_l, \quad (61b)$$

because  $\tilde{A}_l$  commutes with  $\tilde{A}$  for all  $l$  under the stated assumptions. Since the domain of  $\zeta_s$  is the unit sphere  $\mathcal{S} \subset U^\dagger Z \subset \mathfrak{u}(\mathbf{n})$ ,  $A$  is critical point of  $\zeta_s$  if and only if  $d_A \zeta_s(\delta A)$  is zero for all  $\delta A$  perpendicular to  $A$ , i.e. if and only if (61b) is proportional to  $\tilde{A} = \sum \alpha_l \tilde{A}_l$ . So for a given  $s$  there exist many critical points of  $\zeta_s$  among these  $A$  matrices, in particular the cases in which

$\tilde{A} = z|j\rangle\langle k| - \bar{z}|k\rangle\langle j|$  are critical for all  $s$ . If a more detailed critical point analysis could be performed for  $\zeta_s$  with  $s \in [0, \pi/(2\sqrt{2})]$ , it might lead to a confirmation of Conjecture 1.

In addition to these analytic results that hint at the conjecture, more than one million numerical simulations have been performed to test it. In each simulation, random degeneracy structures were sampled for  $\rho$  and  $\mathcal{O}$ , random eigenvalues chosen, a random critical point and a random normal vector selected. Then  $f(s)$  was computed over the interval  $[0, \pi/(2\sqrt{2})]$  and compared to the function  $\beta_{\min}^2 \sin^2(\sqrt{2}s)/2$ . This was done for different system sizes including  $N = 6$ ,  $N = 17$ ,  $N = 25$ , and 1000 trials with  $N = 256$  (which take much longer to run). In every case Conjecture 1 was satisfied. So seemingly, either the conjecture is true, or any counterexamples that exist are hard to find.

## REFERENCES

- [1] Raj Chakrabarti and Herschel Rabitz. Quantum control landscapes. *Int. Rev. Phys. Chem.*, 26(4):671–735, Oct. 2007. doi: 10.1080/01442350701633300. URL <http://www.informaworld.com/10.1080/01442350701633300>.
- [2] Herschel A. Rabitz, Michael M. Hsieh, and Carey M. Rosenthal. Quantum optimally controlled transition landscapes. *Science*, 303:1998–2001, Mar. 26 2004. doi: 10.1126/science.1093649. URL <http://dx.doi.org/10.1126/science.1093649>.
- [3] Herschel A. Rabitz, Tak-San Ho, Michael M. Hsieh, Robert Kosut, and Metin Demiralp. Topology of optimally controlled quantum mechanical transition probability landscapes. *Phys. Rev. A*, 74:012721, 2006. doi: 10.1103/PhysRevA.74.012721. URL <http://link.aps.org/abstract/PRA/v74/e012721>.
- [4] Michael Hsieh, Rebing Wu, Carey Rosenthal, and Herschel Rabitz. Topological and statistical properties of quantum control transition landscapes. *J. Phys. B: At. Mol. Opt. Phys.*, 41(7):074020, 2008. doi: 10.1088/0953-4075/41/7/074020. URL <http://stacks.iop.org/0953-4075/41/i=7/a=074020>.
- [5] Herschel A. Rabitz, Michael M. Hsieh, and Carey M. Rosenthal. Optimal control landscapes for quantum observables. *J. Chem. Phys.*, 124:204107, 2006. doi: 10.1063/1.2198837. URL <http://link.aip.org/link/doi/10.1063/1.2198837>.
- [6] Tak-San Ho and Herschel Rabitz. Why do effective quantum controls appear easy to find? *Journal of Photochemistry and Photobiology*, 180:226–240, 2006. doi: 10.1016/j.jphotochem.2006.03.038. URL <http://dx.doi.org/10.1016/j.jphotochem.2006.03.038>.
- [7] Rebing Wu, Herschel Rabitz, and Michael Hsieh. Characterization of the critical submanifolds in quantum ensemble control landscapes. *J. Phys. A: Math. Theor.*, 41:015006, 2008. doi: 10.1088/1751-8113/41/1/015006. URL <http://stacks.iop.org/1751-8113/41/i=1/a=015006>.
- [8] Herschel A. Rabitz, Michael M. Hsieh, and Carey M. Rosenthal. Landscape for optimal control of quantum-mechanical unitary transformations. *Phys. Rev. A*, 72:052337, 2005. doi: 10.1103/PhysRevA.72.052337. URL <http://link.aps.org/abstract/PRA/v72/e052337>.

- [9] Michael Hsieh and Herschel Rabitz. Optimal control landscape for the generation of unitary transformations. *Phys. Rev. A*, 77:042306, 2008. doi: 10.1103/PhysRevA.77.042306. URL <http://link.aps.org/doi/10.1103/PhysRevA.77.042306>.
- [10] Tak-San Ho, Jason Dominy, and Herschel Rabitz. The landscape of unitary transformations in controlled quantum dynamics. *Phys. Rev. A*, 79:013422, 2009. doi: 10.1103/PhysRevA.79.013422. URL <http://link.aps.org/doi/10.1103/PhysRevA.79.013422>.
- [11] Norman Steenrod. *The Topology of Fibre Bundles*. Princeton University Press, Princeton, 1951.
- [12] Frank W. Warner. *Foundations of Differentiable Manifolds and Lie Groups*. Springer, New York, 1983.
- [13] I. G. Macdonald. The volume of a compact Lie group. *Invent. Math.*, 56(2):93–95, 1980. doi: 10.1007/BF01392542. URL <http://dx.doi.org/10.1007/BF01392542>.
- [14] Y. Hashimoto. On Macdonald’s formula for the volume of a compact Lie group. *Comment. Math. Helv.*, 72(4):660–662, 1997. doi: 10.1007/s000140050040. URL <http://dx.doi.org/10.1007/s000140050040>.
- [15] Hermann Weyl. On the volume of tubes. *Amer. J. Math.*, 61(2):461–472, 1939. URL <http://www.jstor.org/stable/2371513>.
- [16] Alfred Gray. *Tubes*. Birkhäuser, second edition, 2004.
- [17] Manfredo Perdigão do Carmo. *Riemannian Geometry*. Birkhäuser, Boston, 1992.
- [18] James E. Humphreys. *Introduction to Lie Algebras and Representation Theory*. Springer, New York, 1972.
- [19] Brian C. Hall. *Lie Groups, Lie Algebras and Representations: An Elementary Introduction*. Springer, New York, 2003.
- [20] Takashi Sakai. *Riemannian Geometry*, volume 149 of *Translations of Mathematical Monographs*. AMS, Providence, RI, 1996.
- [21] R. F. Snider. Perturbation variation methods for a quantum boltzmann equation. *J. Math. Phys.*, 5:1580, 1964. doi: 10.1063/1.1931191. URL <http://link.aip.org/link/doi/10.1063/1.1931191>.